PHYSICS-BASED THERMAL IMPEDANCE MODELS FOR THE SIMULATION OF
SELF-HEATING IN SEMICONDUCTOR DEVICES AND CIRCUITS

By

JONATHAN SCOTT BRODSKY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1997

This work is dedicated to my parents,

Lawrence and Jeraldine,

my brother Matthew and sister Alexandra.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

PHYSICS-BASED THERMAL IMPEDANCE MODELS FOR THE SIMULATION OF
SELF-HEATING IN SEMICONDUCTOR DEVICES AND CIRCUITS

By

Jonathan Scott Brodsky

August 1997

Chairman: Robert M. Fox
Major Department: Electrical and Computer Engineering

Inherent in the operation of semiconductor devices is self-heating, an
increase in operating temperature due to a device's own power dissipation. The
magnitude of the self-heating effect can be quantified by the value of the thermal
impedance, which describes the dynamic response of the device temperature to
variations in device power. The thermal impedance is determined primarily by
material properties and device structure. The implication of the self-heating effect is
that the change in temperature can alter the operating characteristics of a device,
which in turn, can affect circuit performance.

The primary focus of this dissertation is the development of physics-based
models for the thermal impedances of semiconductor devices. Models for the thermal
impedances of bipolar and field-effect transistors, on both bulk and silicon-on-

insulator (SOI) substrates, are presented. All of the thermal impedance models were derived from the time-dependent heat conduction equation, resulting in compact analytic expressions for the thermal impedances. The physical nature of the thermal impedance models allows them to scale with the device structure and material properties, and they successfully reproduce results from both measurements and three-dimensional finite-element simulations. A circuit model for thermal coupling between transistors in a common substrate is also presented. The coupling model was used in conjunction with the bulk bipolar thermal impedance model to extract a lumped electrothermal model for multiple-emitter bipolar transistors.

The secondary objective of this work is the provision of an approach for incorporating these models into circuit simulators. It has been shown that the thermal impedance models can be represented by thermal equivalent circuits made up of resistors and capacitors, making them suitable for efficient circuit simulation. The computer program TIPP (Thermal Impedance Pre-Processor) is introduced. TIPP was developed to provide circuit simulators with convenient algorithms for generating thermal equivalent circuits. TIPP can calculate the component values for thermal equivalent circuits from either physical models or measured data, and is easily modified to interface with different circuit simulators.

CHAPTER 1
INTRODUCTION

1.1 Self-Heating Effects in Semiconductor Devices

The physical properties of the materials used to fabricate semiconductor transistors depend on temperature. Therefore, the operating characteristics of a transistor (e.g. electrical currents and potentials), which are determined by the material properties, are also temperature dependent. The temperature at which a transistor operates is determined by the temperature of the surrounding environment (referred to as the "local ambient temperature") and the power dissipated in the device (referred to as the "self-heating effect"). Therefore, the time-dependent temperature of a transistor can be expressed as

$$T(t) = T_{amb} + \int_0^t P(t')h_{TH}(t-t')dt', \qquad (1.1)$$

where h_{TH} is the thermal impulse response and P is the instantaneous power. The second term on the right-hand side of (1.1) represents the temperature rise in the device

$$\Delta T(t) = P \otimes h_{TH}, \qquad (1.2)$$

where $\otimes$ is the convolution operator. The temperature rise can also be expressed in

1

the frequency domain as

$$\Delta T(t) = \pounds^{-1}[Z_{TH}(s) \cdot P(s)], \tag{1.3}$$

where $\pounds^{-1}$ represents the inverse Laplace transform and $Z_{TH}(s)$ is the thermal impedance. The thermal impedance of a transistor describes the dynamic response of the device temperature to variations in device power, and is determined primarily by the material properties and the structure of the device. The transient thermal impedance can be defined as

$$Z_{TH}(t) = \pounds^{-1}\left[\frac{1}{s}Z_{TH}(s)\right], \tag{1.4}$$

which represents the normalized thermal step response.

Since the power dissipation in (1.1) is determined by the operating characteristics of a transistor, it depends on temperature such that

$$P = P(T) = I_{dev}(T) \cdot V_{dev}(T), \tag{1.5}$$

where $I_{dev}(T)$ and $V_{dev}(T)$ represent general currents and potentials within a given device, respectively. Consequently, there is feedback between the thermal and electrical operation of the device. Whereas the transistor temperature is usually assumed to be constant, the electrothermal coupling implied by (1.1) and (1.5) shows that the temperature actually varies with the device operation. Thus, to fully characterize the operation of semiconductor transistors, both the electrical and thermal behavior should be determined.

1.1.1 Bipolar Transistors

In the forward-active mode, the operating characteristics of bipolar junction and heterojunction transistors (BJT's and HBT's) are controlled by the injection and diffusion of minority carriers in the base region. For an npn transistor, electrons are injected across the forward-biased base/emitter junction, causing an exponential increase in the minority carriers in the base. The electrons diffuse across the base and are swept into the collector by the reversed-biased base/collector junction. For a fixed base/emitter voltage, assuming negligible recombination in the quasi-neutral base, the collector current can be expressed as

$$I_C(T) \propto n_i^2(T) \cdot \exp\left(\frac{q V_{BE}}{kT}\right), \tag{1.6}$$

where n_i is the intrinsic carrier concentration, q is the electron charge, k is Boltzman's constant and T is temperature. The overall temperature dependence of (1.6) is dominated by the relation between the intrinsic carrier concentration and temperature, given by

$$n_i^2(T) = N_c \cdot N_v \cdot \exp\left(\frac{-E_g}{kT}\right), \tag{1.7}$$

where E_g is the semiconductor band-gap energy and N_c and N_v are the effective density of states in the conduction and valence bands, respectively. The junction voltage is always less than the band-gap and therefore, an increase in temperature causes an exponential increase of minority carriers in the base, resulting in an increase in collector current. Since the collector current is a significant component

of the power dissipation in a BJT, self-heating results in a regenerative feedback between the collector current and the temperature of the device. This positive feedback can lead to the destructive phenomenon of thermal runaway in BJT's [Shu90].

For fixed base current, the collector current can be expressed as

$$I_C(T) = \beta(T) \cdot I_B, \tag{1.8}$$

where $\beta(T)$ is the common-emitter current gain. For moderate injection levels, the current gain can be approximated by the ratio of the electrons injected into the base to the holes back-injected from the base to the emitter. This ratio, and hence $\beta(T)$, are typically high since the emitter usually has a higher doping level than the base. Due to heavy-doping effects in the emitter, the emitter band-gap is typically less than that in the base so that

$$\beta(T) \propto \frac{N_{DE}}{N_{AB}} \exp\left(\frac{-\Delta E_g}{kT}\right), \tag{1.9}$$

where ΔE_g is the band-gap difference between the emitter and base, and N_{DE} and N_{AB} are the doping concentrations in the emitter and base, respectively. As shown by (1.9), the current gain is greater at higher temperatures; consequently, the collector current is, again, an increasing function of temperature. The rate of increase with temperature in this case, however, is not as significant as that for a device biased with a fixed base voltage. Therefore, the self-heating effect is not as substantial in BJT's driven by a fixed base current.

HBT's are bipolar devices that use band-gap engineering in either the emitter or base region to improve the current gain over homojunction BJT's. The resulting band-gap in the emitter is wider than that in the base, so that the potential barrier induced by the band-gap discontinuity effectively impedes the injection of carriers from the base to the emitter. When biased with a fixed base voltage, the temperature dependence of an HBT is similar to that of a standard BJT. However, when an HBT is driven with a fixed base current, the temperature dependence of the collector current is quite different than that of a BJT. While the collector current in this case can still be determined from (1.8), the common-emitter current gain is now expressed as

$$\beta(T) \propto \frac{N_{DE}}{N_{AB}} \exp\left(\frac{\Delta E_g}{kT}\right). \tag{1.10}$$

As a result of the band-gap being wider in the emitter than in the base, the sign of the exponential argument is now positive. Therefore, as opposed to a standard BJT, the current gain and the collector current decrease with increasing temperature. As a result, self-heating in HBT's can lead to the non-catastrophic failure mechanism known as current collapse [Sei93].

To reduce the effects of parasitic resistances and current-crowding, large bipolar devices are commonly fabricated using multiple devices connected in parallel [Shu90]. Multiple-emitter devices, both BJT's and HBT's, are capable of operating at high frequencies under high power densities [Win67, Mar93, Liu95b]. However, multiple-emitter devices suffer from more complex self-heating effects due to the thermal interactions among neighboring devices. The thermal coupling leads to

lateral temperature gradients across the device, resulting in the inner emitters operating at higher temperatures. Due to the positive feedback between junction temperature and junction current, the inner devices carry more current than those at the outer extremes. As the current density in the inner emitters increases, the self-heating effect in these devices accelerates. The premature activation of thermal runaway in BJT's and current collapse in HBT's is attributed to this thermal instability inherent in multiple-emitter devices [Win67, Liu93, Kag94, Lio94, Lio96].

1.1.2 Field-Effect Transistors

For Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFET's) operating in strong inversion, the current characteristics are determined by the drift current of carriers in the inverted channel region. For small drain voltages, a MOSFET operates in the linear region where the carrier velocity depends on the longitudinal electric field in the channel. In the linear region, the drain current can be approximated as

$$I_D(T) \propto \mu(T) \cdot \left(V_{GS} - V_t(T) - \frac{V_{DS}}{2} \right) V_{DS}, \qquad (1.11)$$

where $\mu(T)$ is the carrier mobility, $V_t(T)$ is the threshold voltage, and V_{GS} and V_{DS} are the applied voltages between the gate and source and drain and source, respectively. At higher drain voltages, the electric field at the drain end of the channel is large enough to cause the carrier velocity to saturate. In the saturation

region, the drain current can be expressed as

$$I_D(T) \propto Q_c(V_{GS}, V_{DS}, \mu(T)) \cdot v_{sat}(T), \qquad (1.12)$$

where Q_c is the channel charge and $v_{sat}(T)$ is the saturated carrier velocity.

The overall temperature dependences of (1.11) and (1.12) are dominated by the sensitivity of the carrier mobility to changes in temperature. Due to increased lattice scattering at higher temperatures, mobility decreases as temperature increases. The reduction in mobility leads to a decrease in drain current, which implies that the drain current of a MOSFET is a decreasing function of temperature. At high power dissipation levels, the self-heating effect can cause the drain current to drop below the ambient temperature value. In such cases, the output conductance becomes negative, and the device exhibits a negative dynamic resistance (NDR) [Sha83].

MOSFET's fabricated on silicon-on-insulator (SOI) substrates have temperature dependences that are similar to those of their bulk counterparts, though the effects of self-heating can be enhanced due to the low thermal conductivity of the insulating layers. For non-fully depleted (NFD) SOI MOSFET's, however, floating-body effects further complicate the thermal effects [Wor97]. Impact-ionization-induced floating-body effects are known to cause the kink, or increase in drain current, in NFD SOI MOSFET's. The kink is affected by self-heating in two ways. First, at elevated temperatures, the onset of the impact-ionization is retarded. Second, an increase in recombination in the quasi-neutral body reduces the threshold-voltage shift caused by the impact-ionization. Therefore, in addition to a reduction in drain current due to mobility effects, self-heating also reduces the

current in NFD SOI MOSFET's through temperature-dependent floating-body effects.

1.2 Self-Heating Effects in Semiconductor Circuits

Since the operating characteristics of transistors are affected by temperature, the integrated circuits that depend on these transistors will also be affected by changes in temperature. In modern digital circuits, the high switching speeds of the transistors, the relatively slow time constants associated with the temperature response and the low static power dissipation, all help reduce the instantaneous temperature rise. Consequently, self-heating effects are typically negligible in digital circuits. On the other hand, analog circuit applications commonly have significant power dissipation and can operate at frequencies which are comparable to the thermal time-constants. Therefore, analog circuits are generally more prone to self-heating effects.

1.2.1 Small-Signal Circuit Performance

The effect of self-heating on small-signal BJT characteristics was derived by Mueller and investigated in bipolar circuits by Fox et al. [Mue64, Fox93b]. The two-port small-signal admittance parameters, in the presence of self-heating, were shown to be

$$y_{mn} = \frac{y_{mnE} + D_m Z_{TH} I_m I_n}{1 - D_m Z_{TH} P} \qquad (1.13)$$

where y_{mnE} are the admittance parameters neglecting self-heating, and D_m represents the variation of the current I_m with temperature. The denominator of (1.13) establishes the sensitivity of the admittance parameters to power, and has a significant impact as $D_m Z_{TH} P$ approaches unity. The effect of the denominator generally becomes important only at high power dissipation. However, as the thermal impedance increases (i.e. due to device scaling), the power level that defines the threshold for self-heating effects will decrease. The second term in the numerator of (1.13) shows that the effect of self-heating on the admittance parameters is also proportional to the operating currents. For y_{11} and y_{21} the self-heating term in the numerator is small so that $y_{11} \cong y_{11E}$ and $y_{21} \cong y_{21E}$. However, the effect of self-heating can be substantial in the numerators of y_{12} and y_{22}, even at moderate current levels. The thermal effects on these parameters can result in a coupling between the collector output admittance and the impedance of the base-driving source. Also, as shown in Figure 1.1, there can be a significant reduction in the voltage gain of BJT amplifiers.

The small-signal performance of analog MOSFET circuits can also be affected by self-heating. For moderate power levels, the thermal effects are similar to those in bipolar circuits. However, as mentioned previously, the drain current of a MOSFET decreases with increasing temperature, and significant self-heating can induce NDR. The effect of a negative output conductance can be investigated by examining the voltage gain of a MOSFET amplifier. As shown by Fox and Brodsky [Fox93a], if the devices in the amplifier enter a region of negative output conductance, the gain of the amplifier changes polarity. For an inverting amplifier,
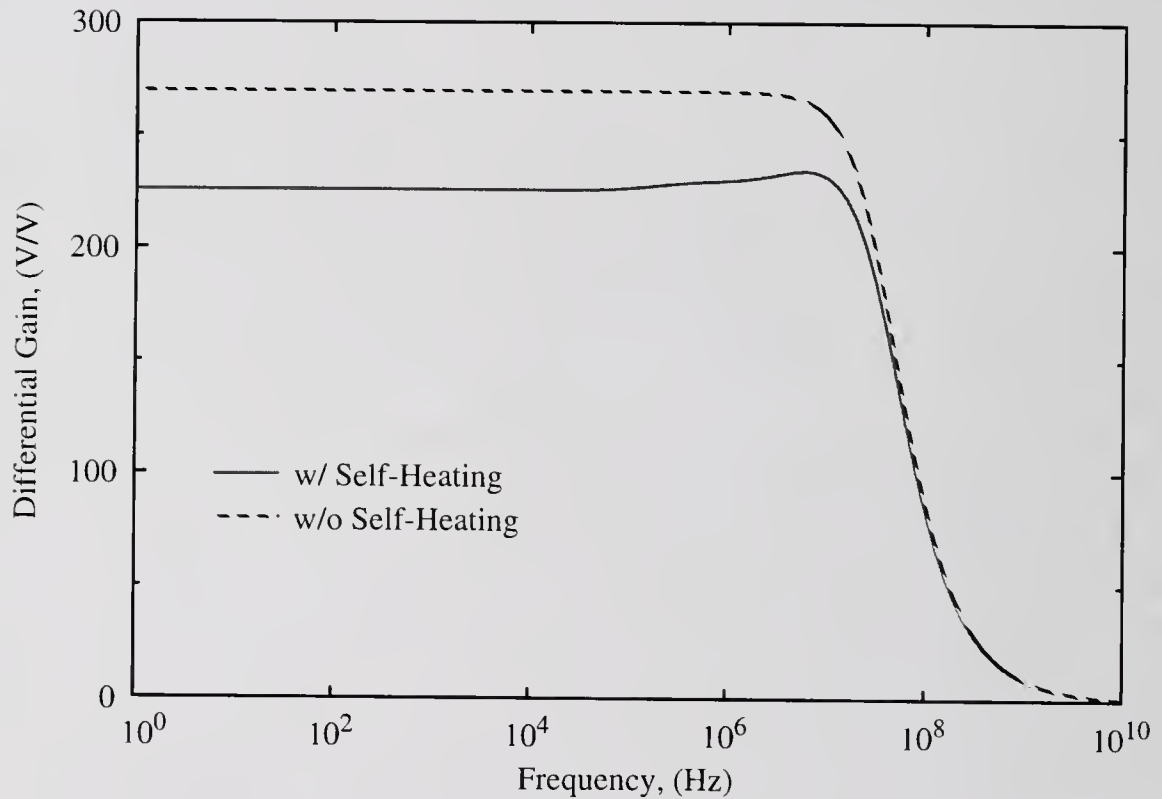
Figure 1.1   The effect of self-heating on the small-signal gain of a BJT differential amplifier. The data was simulated using a version of SPICE, which was modified to account for dynamic variations in temperature [Zwe97], and the thermal impedance model for bipolar transistors presented in Chapter Two.

self-heating effects can therefore cause the gain to become non-inverting, resulting in hysteresis in the amplifier's output characteristics.

1.2.2 Large-Signal Circuit Performance

The effects of self-heating on the large-signal operation of analog bipolar circuits was investigated by Fox et al. [Fox93b]. The types of circuits that are sensitive to thermal effects are typically those that depend on the precise control of BJT characteristics. For example, the mismatch in the reference and output currents of a current mirror can be increased due to self-heating-induced differences in the operating conditions of the transistors. Translinear circuits and band-gap voltage references can also be affected by self-heating due to their strong dependence on the thermal voltage. Thus, neglecting self-heating can result in significant discrepancies between the ideal and actual operation of these types of circuits. The large-signal transient operation of analog circuits is also affected by self-heating. The long time constants of the thermal characteristics can effectively slow down the electrical response of a circuit. Fox et al. showed that the five-percent settling time of a Gilbert multiplier increased by over an order of magnitude due to self-heating [Fox93b]. While the errors caused by self-heating can be reduced by careful circuit design, they can not be completely eliminated.

1.3 Self-Heating Effects in Parameter Extraction

The extractions necessary to determine the behavioral characteristics of a semiconductor transistor are often performed at bias levels that cause moderate- to high-power dissipation. Typically, the parameters that are extracted are assumed to correspond to the ambient temperature at which the measurements are carried out. However, at significant power levels, self-heating will cause a temperature rise in the device. Neglecting the temperature rise that can occur during the measurements can lead to erroneous results [Zwe97]. For example, the Early voltage, V_A, of a BJT is commonly extracted from the slope of the I_C - V_{CE} characteristics in the linear region of operation. If self-heating is significant, the slope of the output curves depends on the source that is driving the base [Fox93b]. Therefore, the exact meaning of the value extracted for V_A would be ambiguous unless the thermal effects were taken into account.

Various methods have been proposed for removing the effects of self-heating from parameter extraction. One approach augments a standard extraction routine with measurements designed to determine the thermal characteristics of the device. The full set of parameters can then be input to a global optimization routine to generate electrical-only parameters that are independent of self-heating [Zwe97]. Other techniques attempt to directly remove the effects of self-heating from the parameter extraction measurements by making the temperature rise in the given device negligible. The temperature rise can be minimized by performing the extractions in low-power regions, or by using complex high-speed measurements [Tu94, Jen95]. Since the device is not allowed to heat, the resulting device parameter

set would be approximately devoid of self-heating effects, and would essentially correspond to the given device operating under isothermal conditions. Consequently, the resulting electrical parameters would only pertain to device operation for low-power or high-speed circuit applications, and would not convey the proper device characteristics for applications that experience substantial self-heating effects. Thus, for a set of electrical-only parameters to correctly represent the characteristics of a device, over a wide range of operating conditions and biases, it should be augmented by additional parameters that describe the thermal attributes of the device.

1.4 The Simulation of Self-Heating Effects

As shown in the previous sections, the operating characteristics of both individual transistors and circuits depend on temperature. Due to self-heating, the effective operating temperature depends on power dissipation and can therefore vary under different operating conditions.

By solving the time-dependent heat conduction equation and energy balance equations for electrons and holes, numerical device simulators can model phenomena associated with dynamic self-heating in individual transistors [Lia94]. While this approach is invaluable for examining the detailed physics that govern the operation of semiconductor devices, it is impractical for simulating all but the simplest of circuits. Therefore, to investigate the effects of dynamic self-heating on a broad range of circuits, a more efficient simulation approach is necessary.

The standard version of most circuit simulators such as Berkeley SPICE [Nag75] and HSPICE [Hsp92], treat temperature as a static global parameter. This has two significant implications. All of the semiconductor devices in a simulation operate at the same temperature, and that temperature remains constant throughout the simulation. Due to these constraints, a circuit simulator may not accurately represent the physical operation of a circuit, where spatial and temporal variations of the temperature can cause each device to operate at its own local temperature. To account for the temperature dependence of a circuit's operation, circuit simulators should be capable of independently tracking the dynamic temperature of each device in the circuit.

A common approach for creating an electrothermal circuit simulator (ETCS) uses the concept of the thermal impedance and the analogy between electrodynamics and heat flow to account for dynamic temperature variations. This approach allows temperature to be represented as an electrical potential and power as an electrical current [Lee96, Zwe97]; therefore, the local operating temperature of a device can be thought of as simply another "bias" condition. To facilitate the temperature "bias" condition, an external node is added to a given compact device model [McA92, Fos95, Lee96, Zwe97]; such a configuration is shown in Figure 1.2. Attached to this node, internal to the device model, is a controlled current source that represents the instantaneous power dissipation. The parameter set for the device model should be modified to include the correct temperature dependences. When the modified device model is used for a circuit simulation, a thermal impedance (and, in some case, a voltage source to represent the reference ambient temperature) can be
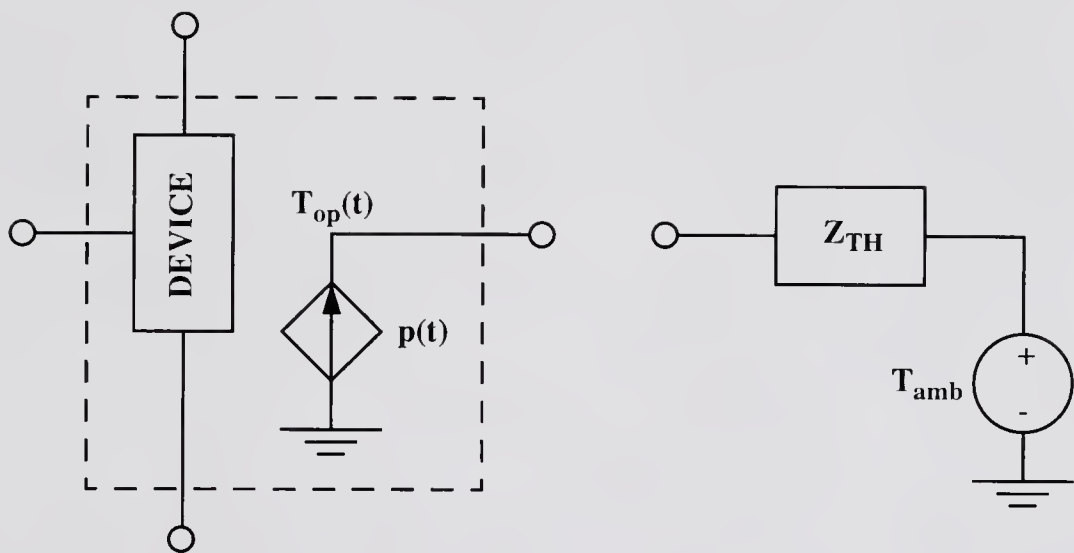
Figure 1.2   A generalized schematic showing a common method for modifying a
compact device model to include temperature as a variable. The dashed
box outlines the new model with the added temperature node; DEVICE
represents the original electrical-only model.

attached to the new external node. Therefore, the voltage generated at this additional node represents the local temperature of the device. The electrical-only device model is first solved at the ambient temperature; this solution results in an initial guess for the device power dissipation. This power is then used to calculate the temperature rise in the device. Once the approximate local operating temperature is calculated, it is used to update the temperature-dependent model parameters, which are used to recalculate the electrical bias potentials and currents of the device model. This procedure is repeated until self-consistent solutions for the temperature and electrical biases are reached. Thus, an effective operating temperature can be independently calculated for each device in a simulation, and that temperature can now vary with the operating point.

1.5 Thermal Equivalent Circuits

The data that quantify the thermal impedance of a transistor are typically in the form of discrete data points for the temperature rise, normalized to a unit-step increase in power dissipation, versus time or frequency. In such a format, the thermal impedance data are not readily accessible by an ETCS. While data in a tabular format can be used without much complexity for DC and AC simulations, an inefficient convolution computation would be required to use the data for transient simulations. Therefore, a representation for the thermal impedance is needed that both accurately models the physical data and can be easily incorporated into an ETCS for efficient DC, AC and transient simulations.

In an ETCS, when a current representing the power dissipation in a transistor is applied to the thermal impedance, Z_{TH}, the resulting voltage represents the temperature rise in that transistor. By invoking the analogy between electrodynamics and heat flow, the thermal impedance can be represented as an electrical impedance. Common representations for the electrical impedance circuits are shown in Figure 1.3. The resistances and capacitances that comprise the impedance effectively represent the lumped three-dimensional thermal resistance and heat capacity of the semiconductor device structure. Therefore, the overall electrical network can be referred to as a thermal equivalent circuit. The values for the individual elements of a thermal equivalent circuit can easily be determined by numerically fitting the circuit to existing thermal impedance data. Thermal equivalent circuits are directly applicable for DC and AC electrothermal simulations since, in such cases, the voltage drop across the network is simply equal to the product of the current and the network resistance or impedance. In addition, such networks inherently provide an efficient method for effecting the necessary transient convolution.

As will be shown in the subsequent chapters of this dissertation, the fundamental nature of heat flow is that of a distributed system. The dynamic temperature rise in a device due to self-heating can occur over three or more decades of time or frequency. A single time constant associated with a simple exponential function can not represent the distributed behavior of self-heating. Consequently, the network response of the single-pole thermal equivalent circuits which have been proposed in previous works [McA92, Bau93, Lee93, Tu94], will not accurately
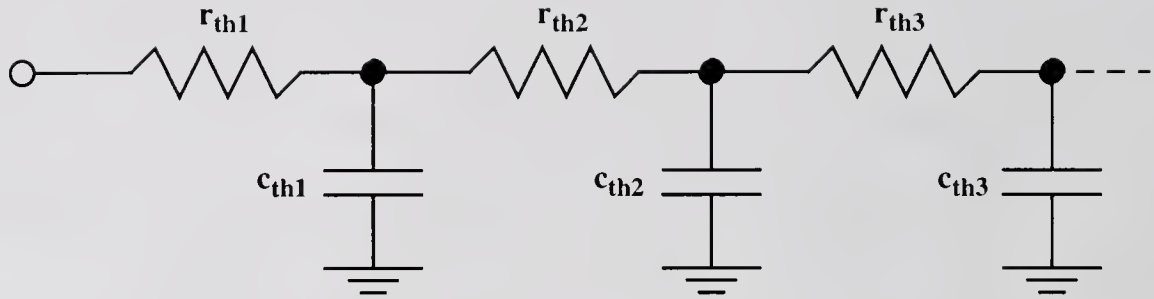
(a)
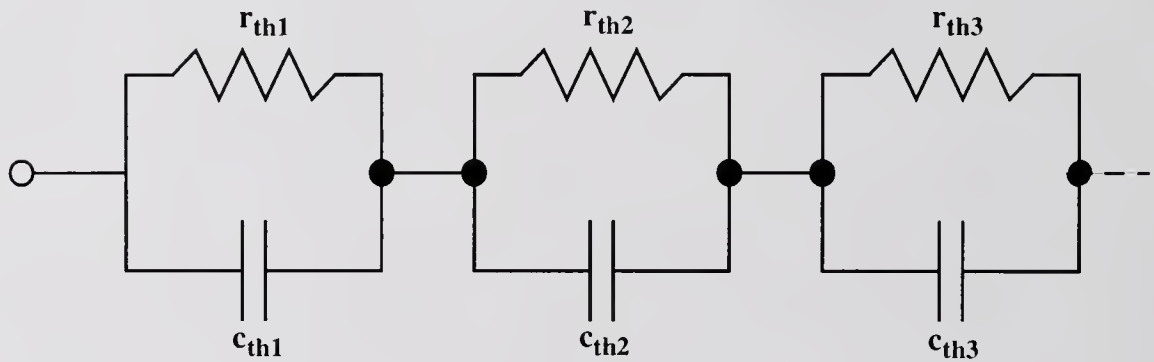
(b)

Figure 1.3   Thermal equivalent circuits used to represent a thermal impedance for circuit simulation: a) Cauer network representation; b) Foster network representation.

model the dynamic thermal impedance. Thermal equivalent circuits consisting of cascaded resistor/capacitor stages, as exemplified in Figure 1.3, effectively provide a distributed network response, and therefore allow a more accurate representation of a dynamic thermal impedance [Bro93].

In the work by Szekely and Van Bien [Sze88], the Foster circuit (Figure 1.3b) was shown to be an invalid representation of a discretized thermal network. This point is valid in the context of numerical simulations (e.g. finite difference or finite element) where the transistor structure is modeled by a distributed thermal network. In that case, the node-to-node capacitances of the Foster network do not have physical meaning and the Cauer network would be the proper physical discretization of the given thermal domain. However, in this dissertation, the thermal equivalent circuit is simply a numerical representation of a thermal impedance, and the validity of its format is moot. Yet, for the purpose of representing a lumped thermal impedance in an ETCS, the Foster network form offers an important advantage over the Cauer form: the time constants associated with a given Foster network are independent of any surrounding circuit elements. This characteristic is beneficial when individual thermal equivalent circuits must be connected to model different components of a transistor structure or the thermal interactions between transistors. Therefore, the Foster network form will be assumed for any thermal equivalent circuits within this dissertation.

1.6 The Need for Physics-Based Thermal Impedance Models

In the previous two sections, the concept of the thermal impedance is adopted to model the temperature rise in a transistor as a function of that device's power dissipation. For the purpose of circuit simulation, the thermal impedance can be represented by a network of resistances and capacitances that effectively represent the lumped thermal characteristics of a transistor. To successfully synthesize a thermal equivalent circuit, tangible data for the thermal impedance are necessary.

One approach to obtain the thermal impedance of a transistor is to extract it from measurements [Lee95, Zwe96]. While this empirical approach provides accurate temperature information, such measurements are somewhat difficult, for several reasons. To begin with, thermal measurements are very time-consuming. The extraction procedure is generally divided into two steps, the first of which dominates the total measurement time. This step is required to calibrate the relation between the temperature and the physical characteristic that is being used to monitor the temperature (e.g. the base current and drain current in bipolar and field-effect transistors, respectively). The calibration is performed at multiple ambient temperatures at DC and is thus limited by the long time constants associated with steady-state heat flow. To make such thermal measurements requires special measurement equipment such as a thermal wafer chuck or oven to accurately control the temperature of the devices being measured. Finally, the results of any such extraction are limited to the specific device being measured. Thus, the entire procedure would have to be repeated for each transistor structure and transistor type of interest.

Another approach, which avoids the inherent complexities of thermal measurements, is to derive the thermal impedance of a transistor from the physical equations that govern the temperature and heat flow in the device. Physical thermal modeling is desirable because it can give the temperature behavior as a function of the device structure and material properties alone; therefore, the effects of device technology scaling on the thermal impedance can be predicted. The requirements of accurate physical modeling (e.g. multi-dimensional numerical simulations) tend to conflict with the needs for simplicity and efficiency in circuit simulation. However, a thermal impedance model does not need to be absolutely accurate to provide reasonable results within an ETCS. Therefore, by using certain heuristic assumptions, compact physical models for the thermal impedance, suitable for efficient simulation, can be derived. It is important, though, that the correlation between the accuracy of the thermal impedance models and the accuracy of the simulated electrical characteristics of a semiconductor device in the presence of self-heating be understood.

The sensitivity of a given electrical parameter, X, of a semiconductor device to the thermal resistance can be defined as

$$S_X^{R_{TH}} \equiv \frac{R_{TH}}{X} \cdot \frac{\partial X}{\partial R_{TH}}. \tag{1.14}$$

As an example, since the output current of a device is very important for characterizing performance, (1.14) can be used to determine the sensitivity of the collector and drain currents of BJT's and MOSFET's, respectively. Using (1.1) in the steady-state limit and (1.6) with (1.14), the sensitivity of the collector current of a

BJT is expressed as

$$S_{I_C}^{R_{TH}} = \frac{q \cdot (V_g - V_{BE}) \cdot R_{TH} \cdot P}{k \cdot (R_{TH} \cdot P + T_0)^2},$$ (1.15)

where V_g is the semiconductor band-gap voltage. A similar expression for the sensitivity of the drain current of a MOSFET can be determined using the current equation given by Fox and Brodsky [Fox93a], which results in

$$S_{I_D}^{R_{TH}} = -\alpha \cdot \left[\frac{R_{TH} \cdot P}{T_0} + 1\right]^{-1} \cdot \frac{R_{TH} \cdot P}{T_0},$$ (1.16)

where α is typically between 1.5 and 1.8 (assuming that the temperature dependence of the drain current is dominated by the temperature sensitivity of the carrier mobility). The expected level of error in simulated output currents can be approximated by the product of the sensitivity and the anticipated error in the thermal impedance model. Therefore, as shown by (1.15) and (1.16), the relation between the accuracy of the thermal impedance models and the accuracy of the calculated electrical parameters depends on the power dissipation and the sensitivity of the electrical parameters to temperature. Consequently, the level of accuracy of a thermal impedance model is more critical for devices with electrical characteristics that are highly sensitive to temperature (e.g. BJT's as opposed to MOSFET's). At low power dissipation levels, where the temperature rise is small compared to the ambient temperature, the error in the thermal impedance model will not directly correspond to the error in the calculated operating temperature. For a temperature rise of twenty degrees, the sensitivities of the BJT collector current ($V_{BE} = 0.8$) and

MOSFET drain current are 0.7 and -0.1, respectively. In such a case, the error in the electrical characteristics will tend to be lower than the error in the thermal impedance. Whereas at large power dissipation levels, the temperature rise can be much larger than the ambient temperature, and the error in the thermal impedance model will directly correspond to the error in the calculated operating temperature (for a temperature rise of one-hundred degrees, the respective BJT and MOSFET current sensitivities are 2 and -0.4); in which case, large errors in the calculated electrical characteristics can result. Figure 1.4 shows an example of BJT characteristics simulated assuming a 20% error in the thermal resistance model; the simulations were performed using the modified version of SPICE created by Lee [Lee96]. The data clearly shows that for larger temperature rises, the error in the calculated current (due to errors in the thermal impedance model) increases.

The motivation behind this dissertation is the development of compact thermal impedance models for semiconductor transistors. These models can provide a reasonably accurate representation of the dynamic temperature response within a device; more importantly, since the models depend mainly on the physical structure of a device, they can correctly anticipate the effects of technology scaling on the thermal behavior. Physics-based thermal impedance models allow an ETCS to predict dynamic self-heating effects in circuits and can also provide more accurate electrical parameter extraction. In addition, when the thermal impedance models are coupled with physics-based compact device models, the combination provides an efficient tool for studying self-heating in semiconductor transistors.
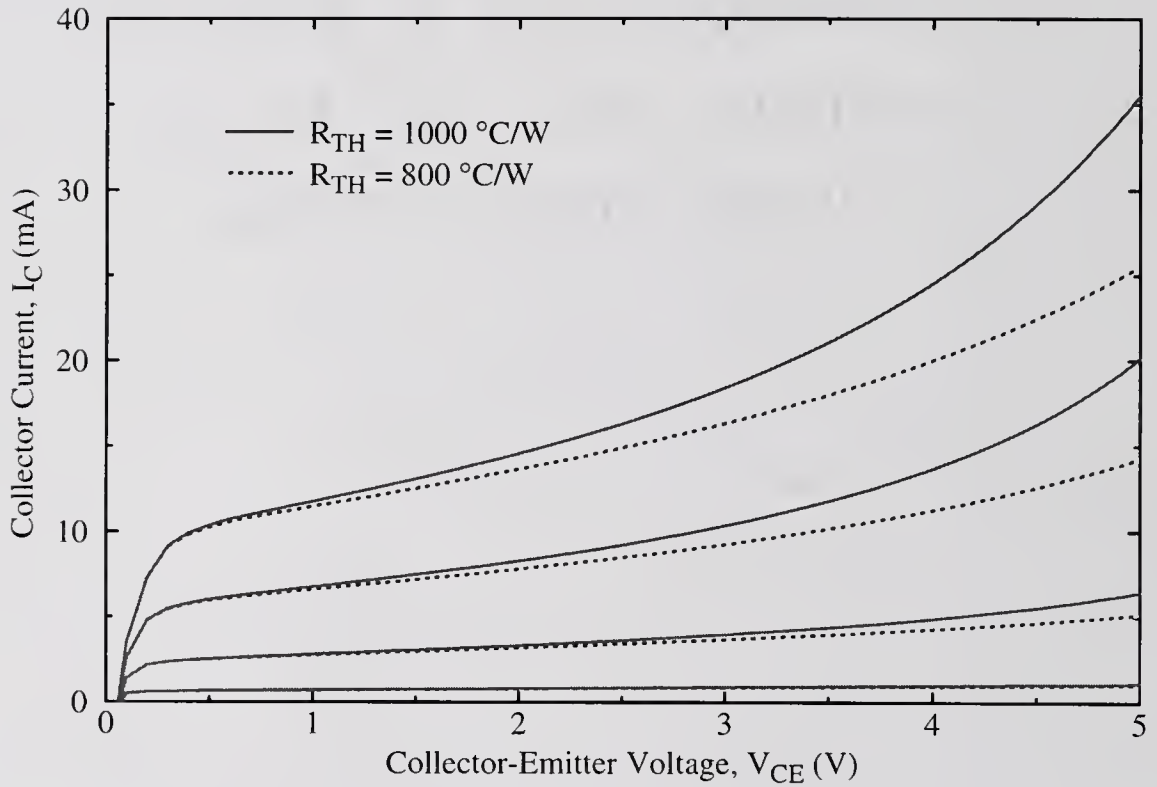
Figure 1.4   Simulated output characteristics of a BJT, assuming that $R_{TH} = 1000\ °C/W$, for $V_{BE} = 0.80,\ 0.85,\ 0.90$ and 0.95 V. The simulations are repeated assuming a 20% error in the thermal resistance model, so that $R_{TH} = 800\ °C/W$.

1.7 Organization

Chapter Two presents a physics-based model for the thermal impedance of bulk junction-isolated bipolar transistors. The model is derived by solving the three-dimensional time-dependent heat conduction equation in the substrate. The ability of the model to represent bulk BJT/HBT's with either LOCOS or trench isolation is investigated. To account for multiple-emitter bipolar transistors, the thermal impedance model is extended to represent multiple heat sources. The accuracy of the model is evaluated using measurements and three-dimensional finite-element simulations.

Chapter Three describes a circuit network for modeling thermal interactions between devices located in the same substrate. The network is developed for the specific application of multiple-emitter bipolar devices, but is shown to be valid for general cross-substrate thermal coupling in circuits. A method for improving the simulation efficiency of a multiple-emitter BJT/HBT electrothermal model, using a lumped thermal impedance model, is presented. The validity of the lumped modeling approach is supported with comparisons to the full electrothermal model.

Chapter Four presents a predictive scalable model for the thermal impedance of BJT's with full dielectric isolation. The model is derived by solving the three-dimensional time-dependent heat conduction equation in the substrate accounting for the buried oxide and trench isolation. In the limit of steady-state heat conduction, the thermal impedance model is simplified, resulting in a closed-form

model of the thermal resistance. The accuracy of both models is evaluated using three-dimensional finite-element simulations and measurements.

Chapter Five describes a physics-based model for the thermal impedance of bulk MOSFET's. The model is derived by solving the three-dimensional time-dependent heat conduction equation in the substrate. The effects of the device interconnects and isolation structures, such as LOCOS and trenches, on the thermal impedance are investigated. The accuracy of the model is evaluated using measurements and three-dimensional finite-element simulations.

Chapter Six presents a predictive scalable model for the thermal impedance of SOI MOSFET's. The model is initially derived for steady-state heat conduction by coupling separate one-dimensional heat conduction analyses in the silicon film and interconnects. The derivation is then carried out for the case of time-dependent heat conduction, resulting in a model for the dynamic thermal impedance. The accuracy of both models is evaluated using three-dimensional finite-element simulations and measurements.

Chapter Seven describes a computer program developed to facilitate thermal modeling in circuit simulation. The program, referred to as the Thermal Impedance Pre-Processor (TIPP), functions as a framework for obtaining the component values of thermal equivalent circuits from the thermal impedance models presented in Chapters Two through Six.

Chapter Eight concludes the dissertation with a summary of the accomplishments of this work and suggestions for future modeling efforts.

CHAPTER 2
A THREE-DIMENSIONAL THERMAL IMPEDANCE MODEL FOR JUNCTION-ISOLATED BIPOLAR TRANSISTORS

2.1 Introduction

The models derived in this chapter provide closed-form physical solutions for predicting the thermal impedances for single- and multiple-emitter bipolar junction (BJT) and heterojunction bipolar (HBT) transistors, based solely on device geometry and material properties. These models can predict both steady-state and dynamic self-heating due to the semiconductor substrate. Previous works in this area provided values for the thermal impedance of BJT's or HBT's, but were either limited by assumptions or relied on non-predictive measurement techniques. For example, the thermal impedance model derived by Fox and Lee [Fox91a] is limited to single-emitter devices. The analyses in other works only provide models for the steady-state thermal resistance [Lio93, Bau94, Daw94, Lio94, Lio96]. Some authors have used measurement techniques to extract either the steady-state thermal resistance or simple one-pole approximations for the thermal impedance [Bau93, Liu93, Daw94, Liu95a, Liu95b]; in either case, the results do not provide a complete picture of self-heating and are not predictive.

The thermal analysis by Joy and Schlig [Joy70] serves as the foundation for deriving of thermal impedance model, and the first part of this chapter re-

examines this work to provide a clear background for modifications made to the model later in the chapter. The thermal impedance model developed by Joy and Schlig was derived for single-emitter, junction-isolated BJT's operating in the forward-active region. A diagram of a typical junction-isolated npn BJT is shown in Figure 2.1. In this chapter, the basic model is modified to account for variations in substrate thickness. The effects of interconnect metallization and different isolation technologies on the thermal impedance, and thus on the performance of the model for advanced device structures, are investigated. The single-emitter thermal impedance model is finally extended to account for BJT/HBT's with multiple emitter fingers.

2.2 Derivation of the Single-Emitter BJT/HBT Thermal Impedance Model

For the derivation of the single-emitter bulk BJT/HBT thermal impedance model, the semiconductor substrate is represented by a homogeneous semi-infinite half-space with an adiabatic top surface (no heat transfer perpendicular to the surface). The back side of the substrate is assumed to be held at a constant temperature, T_0. Since the substrate material is assumed to be homogeneous, the model most directly applies to junction-isolated transistors. The effects of other types of isolation structures used in bulk technologies, such as recessed LOCOS (local oxidation of silicon) or back-filled trenches, on the thermal response are not taken into account. Figure 2.2 illustrates the simplified device geometry assumed for the model derivation; the diagram focuses on the "electrically active" portion of the device that lies directly beneath the emitter stripe, which has a width W and length
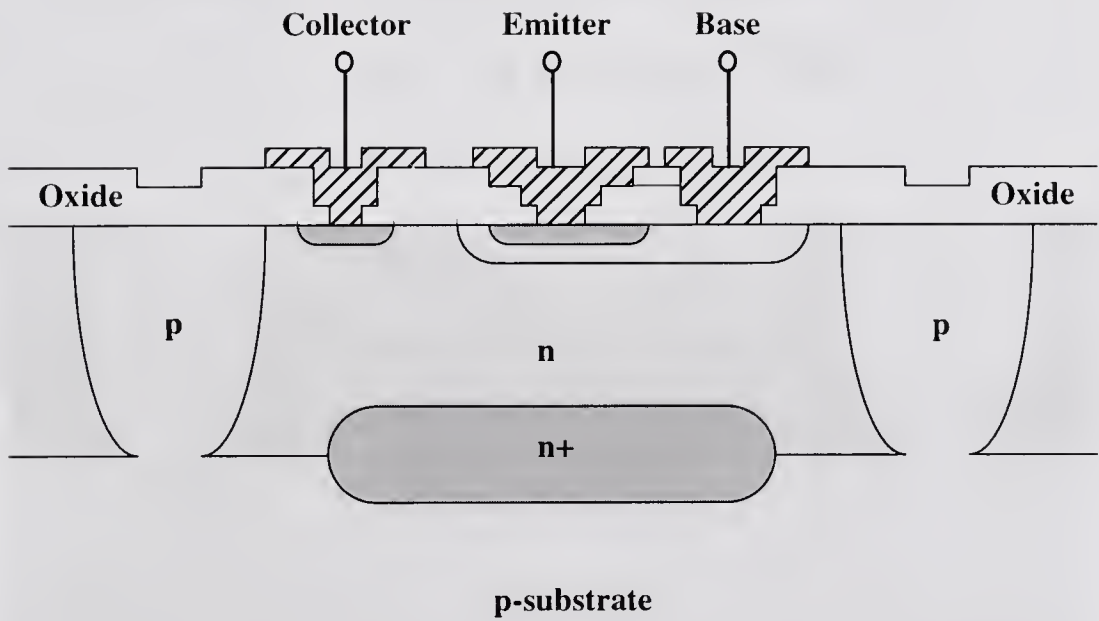
Figure 2.1   Cross-section of a typical junction-isolated bipolar junction transistor (BJT).
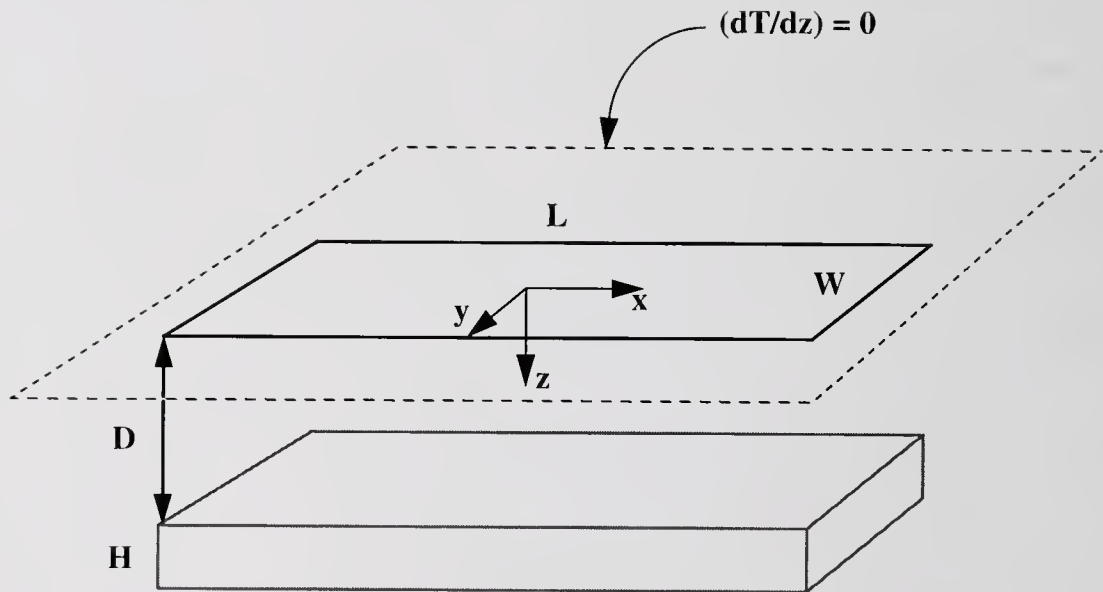
Figure 2.2   The simplified device geometry used to define the solution domain for the bulk, single-emitter BJT/HBT thermal impedance model. The substrate is represented by a semi-infinite half-space with an adiabatic surface (the dotted lines). The emitter stripe has a width W and length L. The heat source (the rectangular volume) is displaced a distance D below the surface of the device, equivalent to the depth of the base/collector junction. The heat source has a thickness H which approximates the base/collector space-charge region (SCR).

L. The imbedded heat source represents the base/collector space-charge region (SCR), which is further represented by a rectangular volume with a thickness, H. The heat generated in this region is assumed to be due to uniform power dissipation. This assumption is reasonable for devices in the forward-active region of operation prior to any high current effects, as the current distribution in the intrinsic device will be approximately uniform. The electric field gradient in the base/collector SCR can also be neglected since it does not greatly affect the thermal impedance model. The heat source is displaced beneath the surface of the substrate by a distance D, assumed to be the depth of the base/collector junction. Thus, any encroachment of the base/collector SCR into the base region is neglected (which is reasonable since the base typically has a higher doping than the collector).

Representing the substrate as a semi-infinite medium presumes that the back-side and the lateral edges do not influence the thermal response of the device. Neglecting the effects of the back side of the substrate on the thermal response is reasonable since a typical wafer is about 1000 times thicker than the heat source. Neglecting the effects of the lateral boundaries requires that the device be located sufficiently far from the substrate edges; the work by Fox et al. [Fox93b] suggests that this assumption is valid for any device that is at least a distance $5\sqrt{W \cdot L}$ from any lateral edge. The surface of the substrate is assumed to be the only boundary that affects the thermal response of the device and it is considered to be adiabatic; thus, conduction through the interconnects and conduction/convection from the surface are neglected. Ignoring thermal energy transport from the substrate surface is supported by the work of Berger and Chai and Goodson et al. [Ber91, Goo95];

however, they were mainly concerned with transport via convection to a surrounding gas (namely air). Nonetheless, for the regions of the device covered by oxide, it is unclear whether there is substantial heat conduction to this overlying oxide. From the analysis of Goodson et al. [Goo95], the device-to-oxide thermal conductance is of the order of $G = 4rk$, which corresponds to an isothermal disk of radius r on the boundary of a semi-infinite medium of thermal conductivity k. Approximating the radius as $\sqrt{(WL)/\pi}$ and using the room-temperature thermal conductivity of SiO_2, the device-to-oxide thermal conductance for a typical device is on the order of 1×10^{-6} (W/°C). Comparably, the device-to-substrate thermal conductance is on the order of 1×10^{-3} (W/°C), showing that the majority of heat will flow through the substrate.

The temperature rise at any point within the device can be described by the nonhomogeneous three-dimensional heat conduction equation

$$\nabla^2 \Delta T(x, y, z, t) + \frac{g(x, y, z, t)}{k} = \frac{1}{\alpha} \frac{\partial \Delta T(x, y, z, t)}{\partial t} \tag{2.1}$$

and the boundary conditions

$$\Delta T(\pm\infty, y, z, t) = 0 \tag{2.2}$$

$$\Delta T(x, \pm\infty, z, t) = 0 \tag{2.3}$$

$$\left. \frac{\partial \Delta T(x, y, z, t)}{\partial z} \right|_{z=0} = 0 \tag{2.4}$$

$$\Delta T(x, y, \infty, t) = 0, \tag{2.5}$$

where ΔT is the temperature rise above the local ambient ($\Delta T = T - T_0$), g is the internal energy generation density, k is the thermal conductivity, α is the thermal diffusivity ($\alpha = k/(\rho \cdot c_p)$ where ρ is the density and c_p is the specific heat) and t is time. Typical values for the material properties are given in Table 2.1.

Table 2.1 Semiconductor material properties

Parameter	Si	GaAs
k (W cm^{-1} K^{-1})	1.412	0.455
ρ (g cm^{-3})	2.328	5.316
c_p (J g^{-1} K^{-1})	0.70	0.35

Source: [Mul77]

Equation (2.1) assumes that the thermal conductivity is independent of temperature and position. Neglecting the temperature dependence of the thermal conductivity is reasonable for a moderate temperature rise, where the temperature rise will vary linearly with power dissipation. However, for large temperature excursions, the value of the thermal conductivity can vary significantly; the thermal conductivities of Si and GaAs will vary from their room-temperature values by more than 20% above 355 and 390 K, respectively [Gao89]. For such large temperature excursions, the linear relation between temperature rise and power will not be valid. However, the temperature dependence of the thermal conductivity can be accounted for by

using the Kirchoff transformation [Joy75], as discussed in Chapter Eight. Neglecting the spatial dependence of the thermal conductivity implies that the effect of dopant atoms is ignored. In the works by Weber and Gmelin and Goodson et al. [Web91, Goo95], the thermal conductivity of doped silicon (up to 1×10^{18} and 1.7×10^{19} dopant atoms cm^{-3}) above 300 K is shown to differ only slightly from that of intrinsic silicon. Since the majority of the substrate is typically low-doped semiconductor material, neglecting the doping effects on the thermal conductivity is reasonable.

With the initial thermal conditions within the substrate specified as

$$\Delta T(x, y, z, 0) = 0, \tag{2.6}$$

the solution to (2.1) can be expressed in the form

$$\Delta T(x, y, z, t) = \frac{\alpha}{k} \int_{t'=0}^{t} dt' \int_{V} G(x, y, z, t | x', y', z', t') g(x', y', z', t') dv' \tag{2.7}$$

where

$$G(x, y, z, t | x', y', z', t') = \frac{1}{8 [\pi \alpha (t - t')]^{3/2}} \exp\left[\frac{-(x - x')^2}{4\alpha(t - t')}\right] \exp\left[\frac{-(y - y')^2}{4\alpha(t - t')}\right]$$

$$\cdot \left\{ \exp\left[\frac{-(z - z')^2}{4\alpha(t - t')}\right] + \exp\left[\frac{-(z + z')^2}{4\alpha(t - t')}\right] \right\} \tag{2.8}$$

is the Green's function for the given boundary-value problem [Ozi93]. Equation (2.8) is the solution to

$$\nabla^2 G + \frac{g_p^i \cdot \delta(x - x')\delta(y - y')\delta(z - z')\delta(t - t')}{k} = \frac{1}{\alpha}\frac{\partial G}{\partial t} \tag{2.9}$$

for the boundary and initial conditions given by (2.2) through (2.6), and physically represents the temperature at point (x, y, z) at time t, due to an instantaneous point source, g_p^i (W s), of unit strength at point (x', y', z') at time t'.

To account for the heat-generation volume $(V = W \cdot L \cdot H)$, (2.8) is substituted into (2.7) and integrated over the base/collector SCR, resulting in

$$
\begin{aligned}
\Delta T(x, y, z, t) = \int_{t' = 0}^{t} \frac{P(t')}{8\rho c V} & \left[\text{erf}\left(\frac{L/2 + x}{\sqrt{4\alpha(t - t')}}\right) + \text{erf}\left(\frac{L/2 - x}{\sqrt{4\alpha(t - t')}}\right)\right] \\
& \cdot \left[\text{erf}\left(\frac{W/2 + y}{\sqrt{4\alpha(t - t')}}\right) + \text{erf}\left(\frac{W/2 - y}{\sqrt{4\alpha(t - t')}}\right)\right] \\
& \cdot \left[\text{erf}\left(\frac{z + D + H}{\sqrt{4\alpha(t - t')}}\right) + \text{erf}\left(\frac{-D - z}{\sqrt{4\alpha(t - t')}}\right) \right. \\
& \left. + \text{erf}\left(\frac{z - D}{\sqrt{4\alpha(t - t')}}\right) + \text{erf}\left(\frac{D + H - z}{\sqrt{4\alpha(t - t')}}\right)\right] dt' \tag{2.10}
\end{aligned}
$$

where $g(x', y', z', t') = g(t') = P(t')/V$, since the power dissipation is assumed to be uniform. Equation (2.10) represents the temperature response at any point in the device at time t due to a change in continuous power dissipation in the base/collector SCR. Assuming a step increase in power at $t' = 0$ ($P(t') = P \cdot U(t)$) and expressing the temperature rise as

$$\Delta T(t) = Z_{TH}(t) \cdot P \tag{2.11}$$

yields the transient thermal impedance

$$Z_{TH}(x, y, z, t) = \int_t \frac{1}{8\rho cV}\left[erf\left(\frac{L/2+x}{\sqrt{4\alpha t}}\right) + erf\left(\frac{L/2-x}{\sqrt{4\alpha t}}\right)\right]$$

$$\cdot \left[erf\left(\frac{W/2+y}{\sqrt{4\alpha t}}\right) + erf\left(\frac{W/2-y}{\sqrt{4\alpha t}}\right)\right]$$

$$\cdot \left[erf\left(\frac{z+D+H}{\sqrt{4\alpha t}}\right) + erf\left(\frac{-D-z}{\sqrt{4\alpha t}}\right)\right.$$

$$\left. + erf\left(\frac{z-D}{\sqrt{4\alpha t}}\right) + erf\left(\frac{D+H-z}{\sqrt{4\alpha t}}\right)\right]dt \qquad (2.12)$$

where the $t \rightarrow \infty$ value of the thermal impedance corresponds to the thermal spreading resistance R_{TH}.

Equation (2.12) represents the temperature rise at any point in the device normalized to a unit-step increase in power dissipation. For circuit simulation, a single temperature is needed to represent the effective operating temperature of the device. Fox and Lee [Fox91b] showed that the thermal impedance model evaluated at a surface corner of the emitter $(x = L/2, y = W/2, z = 0)$ agreed well with measurements of the thermal spreading resistance R_{TH}; substituting these coordinates into (2.12) gives the following expression for the thermal impedance

$$Z_{TH}(t) = \int_t \frac{1}{4\rho cV}erf\left(\frac{L}{\sqrt{4\alpha t}}\right)erf\left(\frac{W}{\sqrt{4\alpha t}}\right)\left[erf\left(\frac{D+H}{\sqrt{4\alpha t}}\right) - erf\left(\frac{D}{\sqrt{4\alpha t}}\right)\right]dt \quad . \qquad (2.13)$$

In this form, the thermal impedance model has four geometric input parameters. Of the four, three (W, L and D) are determined directly by the device layout. However, the fourth parameter, H, depends on the operating bias of the device. The thickness of the base/collector SCR, H, can be estimated using the

depletion approximation; assuming a one-sided step junction with uniform doping on each side gives

$$H = \sqrt{\frac{2 \cdot \varepsilon_{Si} \cdot \varepsilon_0 \cdot (V_R + \Psi_{bi})}{q \cdot N_{epi}}}, \tag{2.14}$$

with

$$\Psi_{bi} = \frac{k_B \cdot T}{q} \cdot \ln\left(\frac{N_b \cdot N_{epi}}{n_i^2}\right), \tag{2.15}$$

where ε_{Si} is the dielectric constant of silicon, ε_0 is the permittivity of free space, V_R is the reverse bias voltage on the base/collector junction, q is electronic charge, N_{epi} is the doping level in the epi-collector, k_B is Boltzman's constant, T is temperature, N_b is the doping level in the base, and n_i is the intrinsic carrier concentration in silicon. Ψ_{bi} is the built-in potential of the base/collector junction. Equation (2.14) shows that the thermal impedance depends on the bias of the base/collector junction, and therefore can change during device operation. However, the square-root dependence of H on the base/collector voltage, is relatively weak. Figure 2.3 illustrates the variation of the modeled thermal resistance with changes in the thickness of the base/collector SCR. The three data points plotted for each simulated device correspond to reverse bias base/collector voltages of 5, 10 and 20 V. The largest variation is observed for the smallest device, which shows a 25% change in its thermal resistance going from $V_R = 0$ V to 20 V. The larger devices show a weaker dependence on H and have no more than a 15% change in thermal resistance
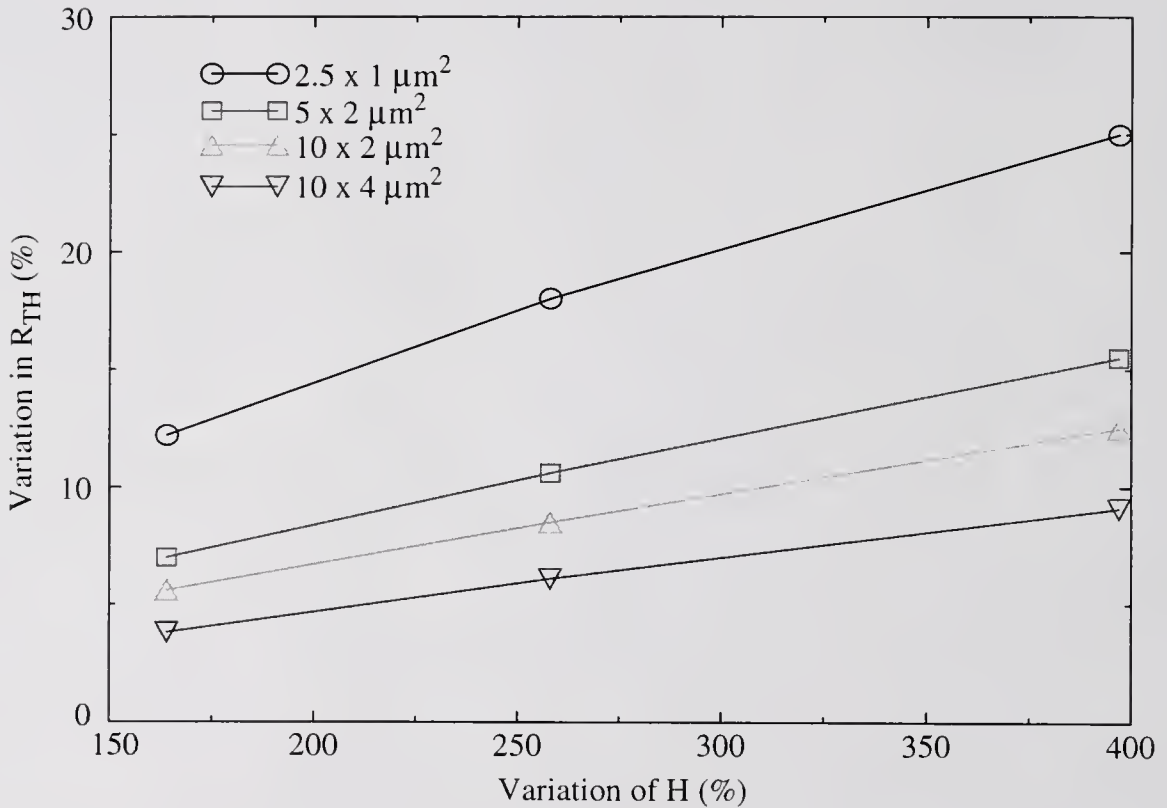
Figure 2.3   Simulations showing the effect of variations in the thickness of the base/collector space-charge region on the thermal impedance model (evaluated at steady-state) for different geometry BJTs. For each device, D = 0.35 μm, N_{epi} = 1 x 10^{16} cm^{-3} and N_b = 1.5 x 10^{18} cm^{-3}. The y-axis corresponds to the variation between the model evaluated at V_R = 0 V and the model evaluated at V_R equal to 5, 10 and 20 V.

going from $V_R = 0$ V to 20 V. However, at high base/collector biases, the maximum value of H becomes effectively independent of bias. As shown in Figure 2.1, typical bipolar technologies use a heavily-doped buried layer to reduce collector resistance. At high base/collector biases, the low-doped epi-collector region depletes down to the buried layer; consequently, the maximum value of H should be properly limited to the thickness of the epi-collector.

2.2.1 Modification for Finite Wafer Thickness

As previously derived, the thermal impedance model for single-emitter bulk BJT/HBT's represents the substrate as a semi-infinite half-space. This representation assumes that the back side of the substrate does not affect the thermal response of the device. In general, this is reasonable since the base-collector junction is usually within 1 μm of the substrate surface, and a typical wafer is between 350 to 800 μm thick. However, wafers are commonly back-lapped to improve thermal performance, and substrate thicknesses of 75, 80 and 100 μm have been reported by a number of authors [Kag94, Mar93, Liu95a]. As the wafer thickness is reduced, the substrate can no longer be approximated by a semi-infinite medium and the effects of the back-side boundary must be taken into account.

The three-dimensional Green's function in the rectangular coordinate system can be represented by the product of three one-dimensional Green's functions

$$G(x, y, z, t | x', y', z', t') = G_x(x, t | x', t') \cdot G_y(y, t | y', t') \cdot G_z(z, t | z', t'). \qquad (2.16)$$

The lateral boundaries are still assumed to extend infinitely and their effects on the thermal response are neglected; thus, the Green's function solutions in both the x and y directions remain unchanged. In the z-direction, however, the substrate is now assumed to have a finite thickness D_{sub}. The top surface of the substrate is still assumed to be adiabatic. The bottom surface of the substrate is assumed to be at a constant temperature, $T(D_{sub}) = T_0$, so that the temperature rise at this surface is defined by $\Delta T(D_{sub}) = T(D_{sub}) - T_0 = 0$. These boundary conditions define the new Green's function for the z-direction, which is given by

$$G_z(z, t|z', t') = \sum_{p = 1}^{\infty} \exp[-\alpha \cdot \eta_p^2(t - t')] \cdot \frac{2}{D_{sub}} \cdot \cos(\eta_p z) \cdot \cos(\eta_p z') \qquad (2.17)$$

where η_p is the set of eigenvalues for the boundary-value problem and are given by the positive roots of

$$\cos(\eta_p D_{sub}) = 0. \qquad (2.18)$$

Equation (2.18) is solved when the argument of the cosine equates to odd multiples of $\pi/2$. Using equations (2.17), (2.16), (2.11) and (2.7), and then integrating over the base/collector SCR, assuming a unit-step increase in power at $t' = 0$, gives the following expression for the thermal impedance at any point in the device

$$Z_{TH}(x, y, z, t) = \int_t \frac{dt}{4\rho c V}\left[\text{erf}\left(\frac{L/2 + x}{\sqrt{4\alpha t}}\right) + \text{erf}\left(\frac{L/2 - x}{\sqrt{4\alpha t}}\right)\right]$$

$$\cdot \left[\text{erf}\left(\frac{W/2 + y}{\sqrt{4\alpha t}}\right) + \text{erf}\left(\frac{W/2 - y}{\sqrt{4\alpha t}}\right)\right]$$

$$\cdot \sum_{p = 1}^{\infty} \frac{2\cos(\eta_p z)\exp(-\alpha\eta_p^2 t)}{\eta_p D_{sub}}$$

$$\cdot \{\sin[\eta_p(D + H)] - \sin[\eta_p D]\} \qquad (2.19)$$

where $V = W \cdot L \cdot H$ and

$$\eta_p = \frac{(2p - 1)\pi}{2D_{sub}} . \qquad (2.20)$$

Evaluating (2.19) at the coordinates ($x = L/2, y = W/2, z = 0$) to give a single effective operating temperature, results in the following expression for the thermal impedance

$$Z_{TH}(t) = \int_t \frac{dt}{4\rho c V}\text{erf}\left(\frac{L}{\sqrt{4\alpha t}}\right)\text{erf}\left(\frac{W}{\sqrt{4\alpha t}}\right)$$

$$\cdot \sum_{p = 1}^{\infty} \frac{2\exp(-\alpha\eta_p^2 t)}{\eta_p D_{sub}}\{\sin[\eta_p(D + H)] - \sin[\eta_p D]\} . \qquad (2.21)$$

Since (2.21) is derived from the physical heat conduction equation, it can be used to anticipate the effects of substrate scaling on the thermal impedance. Figure 2.4 illustrates equation (2.21) evaluated at various values for D_{sub}. The results show that the thermal resistance decreases as the substrate thickness is reduced, which agrees with the trend predicted by Hattori et al. using a three-dimensional numerical simulator [Hat95]. Figure 2.4 also shows that (2.13) provides
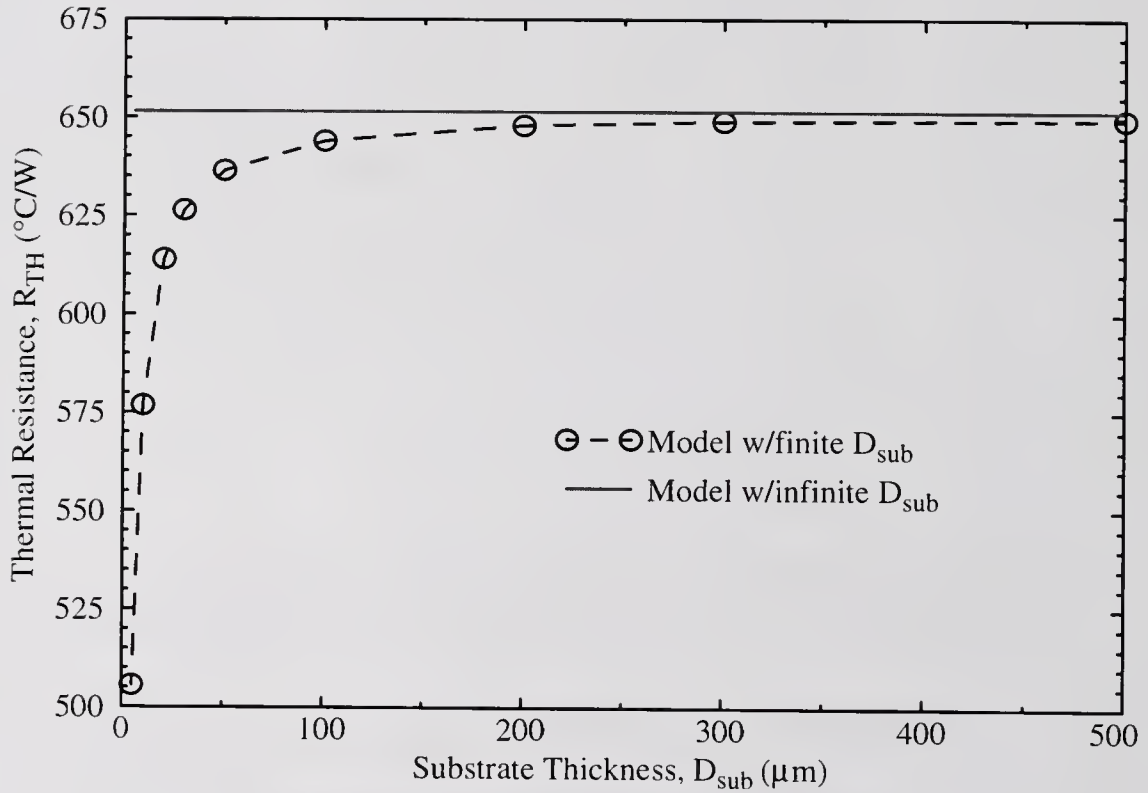
Figure 2.4   Simulations showing the effect of substrate thickness on the thermal impedance model (evaluated at steady-state). The model accounting for finite substrate thickness is compared to the model assuming infinite substrate thickness. The device specifications are L = 4 μm, W = 1 μm, D = 0.35 μm, H = 0.35 μm.

an accurate prediction of the thermal resistance over most of the range of substrate thicknesses; only when the substrate thickness is significantly reduced ($< 100 \, \mu m$), is there a large deviation between the two models.

2.2.2 Effects of Interconnect Metallization on the Thermal Impedance

For the derivation of the BJT/HBT thermal impedance model, the surface of the substrate is assumed to be adiabatic. In actual devices, portions of the base, collector and emitter regions are in direct contact with the metallization used to electrically connect different devices on a chip. Since the metallization typically has a high thermal conductivity, it is possible that the heat conduction via the interconnects significantly influences the thermal response of a device. Therefore, the validity of such an assumption should be investigated.

Three-dimensional (3-D) finite-element (FE) thermal simulations of a bipolar transistor, using the ANSYS software package [Ans96], were performed to examine the effects of the interconnect metallization on the thermal impedance. To simplify the FE model, the device was considered to be symmetric in both lateral directions; therefore, only one quarter of the device was simulated. The bottom and exterior sides of the substrate were assumed to be at a fixed ambient temperature. The top and side surfaces of the interconnects, as well as the top surface of the inter-layer dielectric, were assumed to be adiabatic. The FE simulations tend to overpredict the heat conduction through the interconnects since any contact resistances at the material interfaces were neglected. The assumed symmetry of the device implies that the base and collector metallization are equidistant from the emitter. In typical

devices, the collector contact is offset a greater distance from the emitter than the base contact. Typical ranges for these offsets are 0.5 to 10 µm between the base and emitter contacts, and 2.5 to 25 µm between the collector and emitter contacts [Gra93]. While the FE model does not exactly represent any actual device structure, it can provide an estimate for the significance of the heat flow through the interconnects as a function of their distance from the active device.

Figure 2.5a shows the FE model for bipolar devices with full metallization. Steady-state thermal simulations were run for various interconnect spacings; this spacing corresponds to the edge-to-edge distance between the emitter and base/collector interconnects. Simulations were also run of the same structure with the base/collector interconnect removed. The results of the two groups of simulations were compared to determine the effect of the base and collector interconnects on the thermal resistance. Figure 2.5b shows the results of the comparison between the FE simulations. The data clearly shows that the effect of the base/collector interconnect metallization is small and decreases as the interconnects are moved away from the active device area. The collector interconnect has less of an effect on the thermal impedance than the base interconnect, due to the larger distance between the collector contact and the active device. In any case, the effect of either the base or collector interconnect should be negligible compared to the influence of the emitter metallization.

To determine the extent of the effect of the emitter interconnect on the thermal impedance, steady-state thermal simulations were run for different devices with only the emitter metallization in contact with the device. Figure 2.6a shows the
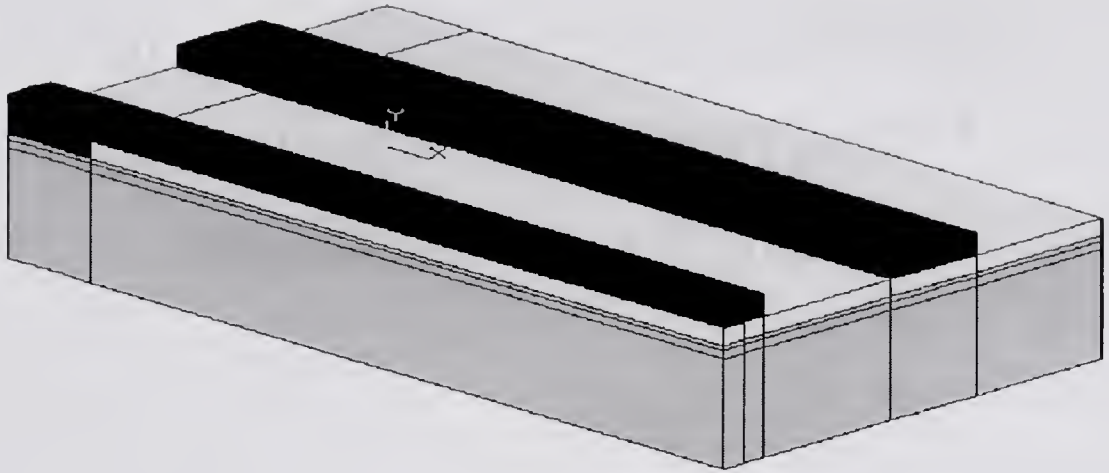
(a)
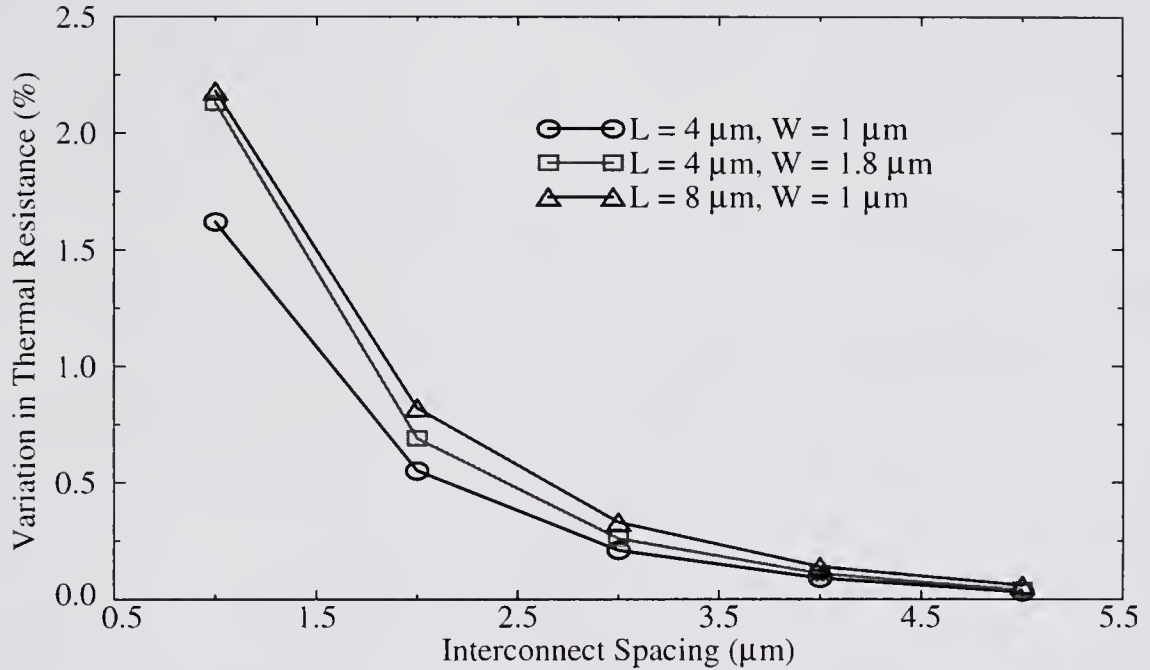
(b)

Figure 2.5   ANSYS simulations showing the effect of emitter, base and collector interconnects on the thermal resistance. The device specifications are $D = 0.2\ \mu m$ and $H = 0.35\ \mu m$, the interconnect width $W_{met} = 2\ \mu m$ and thickness $d_{met} = 0.9\ \mu m$, and the inter-layer dielectric thickness $d_{ox} = 0.7\ \mu m$: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for emitter, base and collector interconnects and the thermal resistance accounting for only the emitter interconnect, plotted as a function of the spacing between emitter and base/collector interconnects.

(a)

(b)
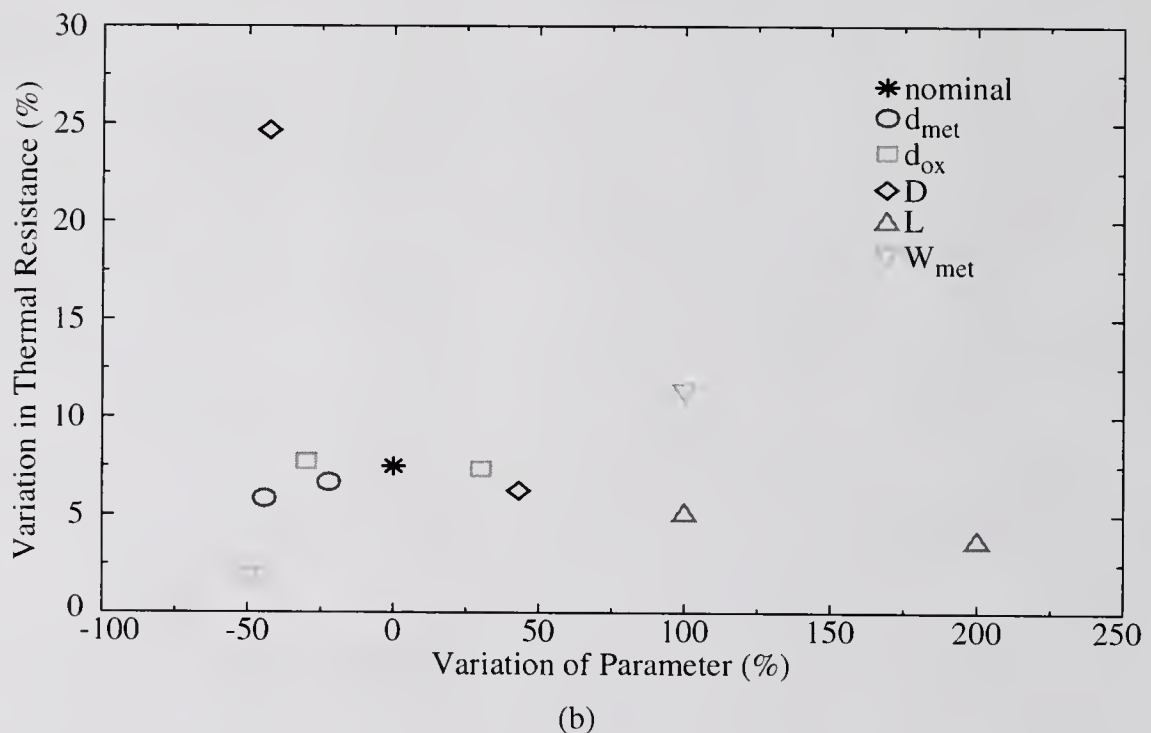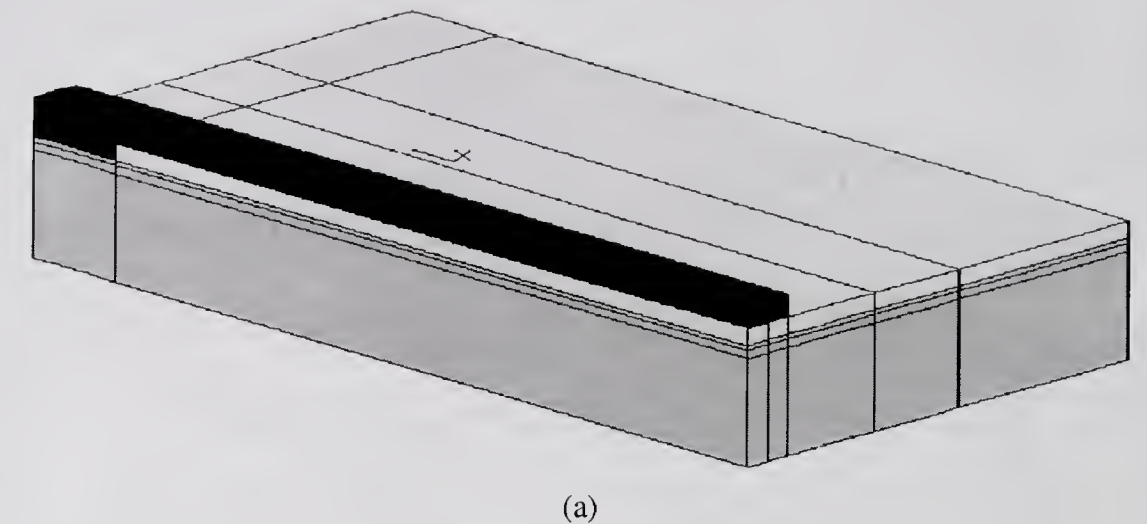
Figure 2.6   ANSYS simulations showing the effect of the emitter interconnect on the thermal resistance for variations in different technology characteristics. The specifications for the nominal device are L = 4 μm, W = 1 μm, D = 0.35 μm, H = 0.35 μm, d_{met} = 0.9 μm, d_{ox} = 0.9 μm and W_{met} = 2 μm: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for the emitter interconnect and the thermal resistance neglecting the emitter interconnect. The variation plotted on the x-axis corresponds to the deviation of each structure parameter from its nominal value.

FE model for devices with only the emitter metallization. The same devices were also simulated with the emitter interconnect removed. The simulations were performed by independently varying each structure parameter of the FE device model. Reasonable values were chosen for each parameter to represent a nominal device design; each parameter was varied about the nominal value to represent a reasonable range of technology scaling. Figure 2.6b shows the results of the FE simulations. The emitter metallization becomes a more effective path for heat evacuation as the thickness, d_{met}, and the width, W_{met}, of the interconnect increase and as the thickness, d_{ox}, of the dielectric layer between the substrate and the interconnect decreases. The data also show that the effect of the emitter interconnect increases as the depth of the base/collector junction is decreased (the heat source is moved closer to the surface) and as the length of the emitter is decreased.

Transient thermal simulations of the FE model in Figure 2.6a were used to examine the effect of the emitter interconnect on the transient thermal response. ANSYS was used to simulate the structure with and without the emitter interconnect in contact with the device; the results are shown in Figure 2.7. The thermal responses for the device with and without the interconnect match until significant heat reaches the surface of the device. The time for heat to reach the surface of the device can be approximated as the square of the distance D divided by the thermal diffusivity of the substrate material. The resulting time is approximately 0.44 nanoseconds, which agrees with the FE simulations. The adiabatic boundary condition of the device without the interconnect predicts a larger response since the heat is completely reflected once it reaches the surface. The device with the emitter metallization, which
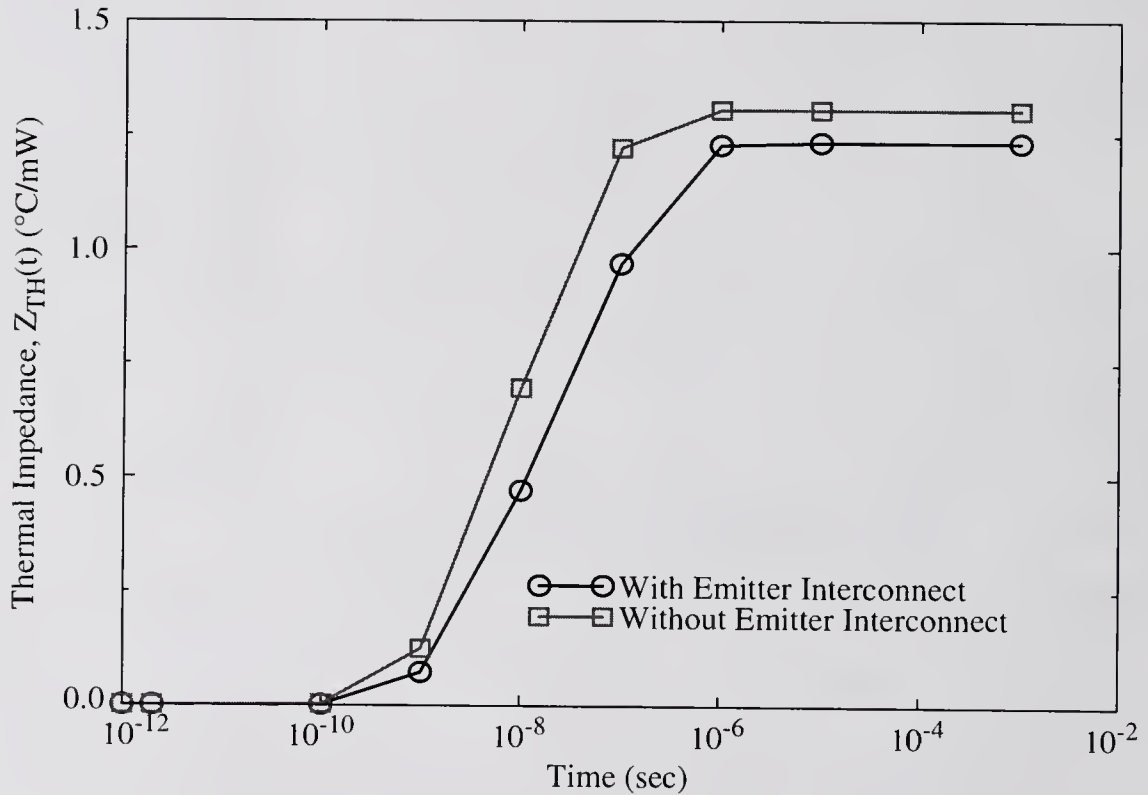
Figure 2.7   ANSYS simulations showing the effect of the emitter interconnect on the transient thermal impedance. The specifications for the device structure are L = 2 μm, W = 1 μm, D = 0.2 μm, H = 0.35 μm, d_{met} = 0.9 μm, d_{ox} = 0.7 μm and W_{met} = 2 μm.

acts as a separate heat sink path, has a reduced thermal impedance and an effectively slower thermal response.

Based on the results of the 3-D FE simulations, neglecting the base and collector interconnect metallization in the model derivation is reasonable since it only slightly affects the thermal impedance of a device. The emitter interconnect, however, has a greater influence on both the steady-state and transient thermal responses. The effect on the thermal resistance will be more significant for devices with small-geometry emitters and shallow base regions, where the transient thermal response will mainly be affected for large devices with substantial contact structures. In either case, equations (2.13) and (2.21) will tend to over-predict both the steady-state thermal resistance and the transient rise of the thermal impedance.

2.2.3 A Model for the Thermal Impedance of the Emitter Interconnect

As shown in the previous section, the assumption that the top surface of the device is adiabatic neglects heat flow in the emitter interconnect and results in a thermal impedance model that over-estimates the transient temperature rise in a bipolar device. To model the effects of the emitter interconnect on the overall thermal impedance, both the thermal resistance and thermal capacitance of the metallization need to be considered.

The thermal resistance of the emitter metallization is derived by assuming that the interconnect can be represented by a one-dimensional cooling fin, so that the temperature rise at any point x_{met} along the interconnect, $\Delta T_{met}(x_{met}) = T_{met}(x_{met}) - T_0$, can be approximated by

$$\frac{\partial^2 \Delta T_{met}}{\partial x_{met}^2} - m_{met}^2 \Delta T_{met} = 0. \tag{2.22}$$

The second term on the left-hand side of (2.22) accounts for heat conduction through the underlying oxide as the heat travels along the interconnect, where

$$\frac{1}{m_{met}} = \sqrt{\frac{k_{met} d_{met}}{h_{met}}} \tag{2.23}$$

is the characteristic thermal length in the interconnect and

$$h_{met} = \frac{k_{ox}}{d_{ox}} \tag{2.24}$$

is the heat transfer coefficient from the interconnect to the substrate. The material properties for the emitter interconnect are given in Table 2.2.

Table 2.2 Emitter interconnect[‡] material properties

Property	Definition	Value
k_{met}	Thermal conductivity	2.39 (W cm^{-1} K^{-1})
ρ_{met}	Density	2.7 (g cm^{-3})
c_{pmet}	Specific Heat	0.9 (J g^{-1} K^{-1})

Source: [Ozi93]
[‡] Assumed to be aluminum

Approximating the temperature in the interconnect using a one-dimensional equation implies that the temperature gradients in the vertical and lateral directions within the emitter interconnect are negligible. The validity of such an assumption can be evaluated using the Biot number, which corresponds to the ratio of the internal and external thermal resistances of a given object [Ozi93]. If the Biot number for the interconnect is much less than unity, then the interconnect can be approximated as a one-dimensional thermal medium. The vertical and lateral Biot numbers for the emitter interconnect are given by $B_{Vmet} = h_{met}d_{met}/k_{met}$ and $B_{Lmet} = h_{met}W_{met}^2/(k_{met}d_{met})$, respectively. For most practical metallization geometries, the Biot numbers are much less than one and the cooling-fin model is an accurate representation of the emitter interconnect.

Assuming that the temperature rise in the interconnect at the emitter contact is equal to the effective operating temperature of the device, and that the temperature rise approaches zero far from the contact, the thermal resistance of the emitter interconnect can be expressed as

$$R_{THmet} = [k_{met}m_{met}W_{met}d_{met}]^{-1}. \qquad (2.25)$$

The thermal capacitance of the emitter metallization can be approximated as

$$C_{THmet} = \rho_{met}c_{pmet}V_{met}. \qquad (2.26)$$

The volume of the metallization is $V_{met} = W \cdot L \cdot \delta_{met}$, where δ_{met} represents the effective length of the interconnect structure. The parameter δ_{met} should be evaluated to include the volume of the contact and interconnect metallization but can

also be extracted from transient thermal measurements or numerical simulations. Once the thermal resistance and thermal capacitance have been calculated, the transient thermal impedance of the emitter interconnect can be approximated by

$$Z_{THmet}(t) = R_{THmet}\left[1 - \exp\left(\frac{-t}{\tau_{met}}\right)\right], \tag{2.27}$$

where $\tau_{met} = R_{THmet}C_{THmet}$. The overall thermal impedance of a bipolar device can now be represented by the parallel combination of two thermal impedances, such that effectively

$$Z_{TH}(s) = \frac{Z_{THdev}(s) \cdot Z_{THmet}(s)}{Z_{THdev}(s) + Z_{THmet}(s)}, \tag{2.28}$$

where $Z_{THdev}(s)$ is determined from the transient thermal impedance given by either (2.13) or (2.21).

2.2.4 Effects of Isolation Structures on the Thermal Impedance

While junction-isolated technologies are still used, the drive to increase packing density, improve lateral isolation and increase device operating speeds has led to the development of newer isolation technologies for VLSI bipolar applications. The advanced isolation technologies typically used in bulk bipolar fabrication are recessed LOCOS (local oxidation of silicon) and U-groove [Wol90, Gra93]; Figure 2.8 illustrates examples of bipolar devices fabricated with these isolation techniques. Since advanced isolation structures typically use low conductivity
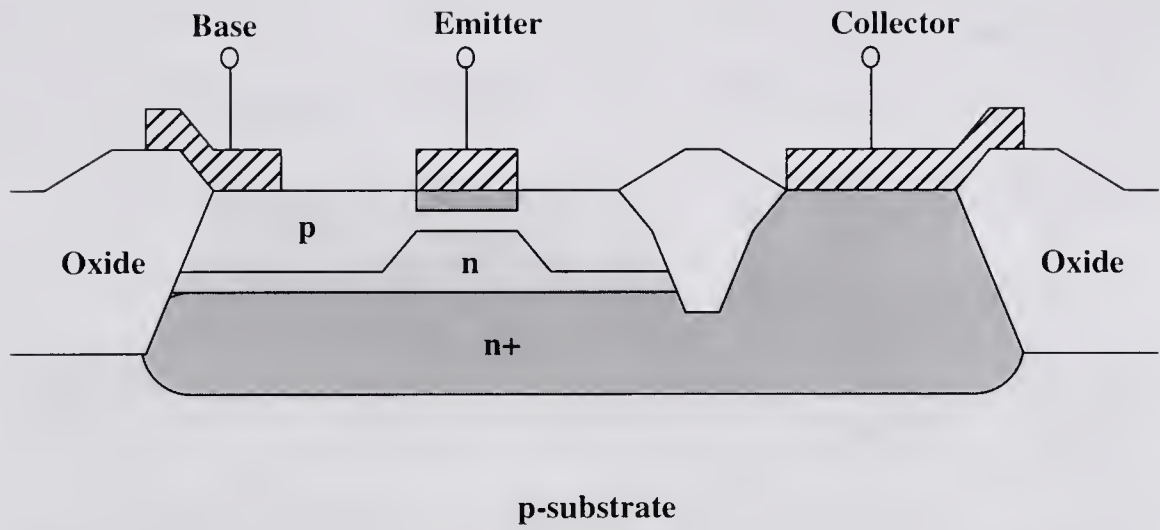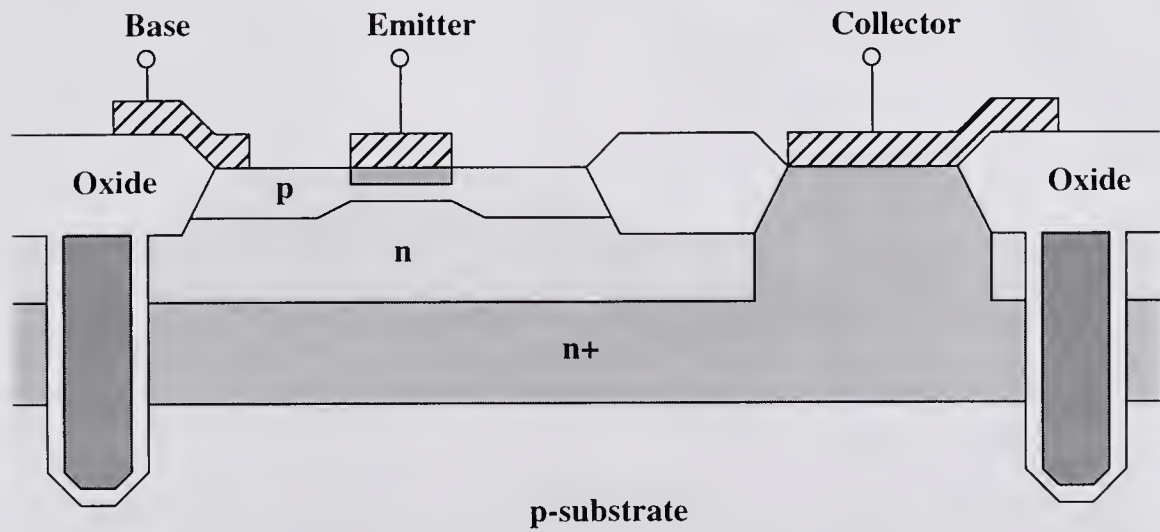
Figure 2.8   Cross-sections of typical BJT's fabricated with advanced isolation technologies: a) Recessed LOCOS; b) U-groove isolation.

materials like SiO_2, the thermal impedance of a device using such isolation tends to be higher than that of its junction-isolated counterpart. The bulk BJT/HBT thermal impedance model treats the substrate as a homogeneous material; therefore, it is unclear whether the thermal impedance model is applicable to devices which are fabricated with advanced isolation.

Three-dimensional (3-D) finite-element (FE) thermal simulations, using ANSYS, were performed to examine the effects of advanced isolation structures on the thermal impedance of bipolar transistors. Two FE models were developed to separately investigate the effects of recessed LOCOS and U-groove isolation. To simplify the FE models, the device was considered to be symmetric in both lateral directions, so that only one quarter of the device was simulated. The bottom and exterior sides of the substrate were assumed to be at a fixed ambient temperature. The top surface of the device was assumed to be adiabatic. Due to the assumed symmetry, the active device region was surrounded on all sides by the isolation structure, which was at uniform distance from each side of the emitter. As shown by the illustrations in Figure 2.8, the distance between the isolation structure and the intrinsic device is not uniform on all sides of the device. Typical values for the distance between the emitter and the isolation--for the portions of the isolation structure immediately surrounding the emitter--are in the range from 0.3 μm to 0.8 μm for advanced bipolar devices [Del91, Klo93, Yam93, Pru94]. The distance between the emitter and the portion of the isolation structure on the far side of the collector contact is generally larger, typically two to four microns [Del91, Klo93]. For both FE models, the side-walls of the isolation structures were perpendicular to the top surface of the substrate.

For typical U-groove isolation, the trench is created by anisotropic etching and the side-walls are nearly perpendicular to the surface. However, actual LOCOS structures have tapered edges (see Figure 2.8) that get progressively thinner toward the active device. To determine the implications of the model's non-physicality, two-dimensional (2-D) FE simulations were run for various angles (30 to 90 degrees) between the substrate surface and the side-wall of the isolation. The simulations showed that the temperature rise increased as the angle increased; thus, the 3-D LOCOS model should show a larger effect than that of an actual isolation structure. While the finite-element models do not truly represent the physical device layout, they allow an order-of-magnitude estimate for the effects of the isolation structures on the thermal impedance.

Steady-state thermal simulations were run for various device-isolation spacings, corresponding to the edge-to-edge distance between the emitter and the isolation structure. For the U-groove isolation model, this spacing is the distance between the emitter and the edge of the surface LOCOS; the actual trench is assumed to be an additional 0.5 μm away from the edge of the LOCOS [Del91, Yam93]. Simulations were also run for the same devices with the isolation structures removed.

LOCOS isolation is formed by selectively oxidizing regions of the semiconductor substrate in a dry or wet oxygen-rich ambient. For bipolar technologies, the resulting SiO_2 structures are typically no more than one micron thick, since the growth of thicker oxides is impractical [Wol90, Gra93]. Figure 2.9a illustrates the FE model for bipolar devices with recessed LOCOS isolation. The oxide was assumed to be fully recessed beneath the top surface of the substrate and
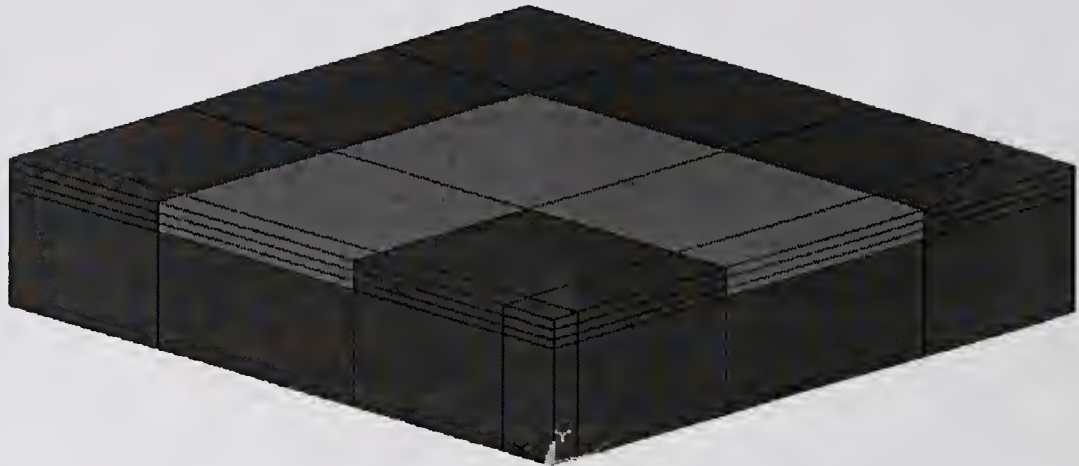
(a)
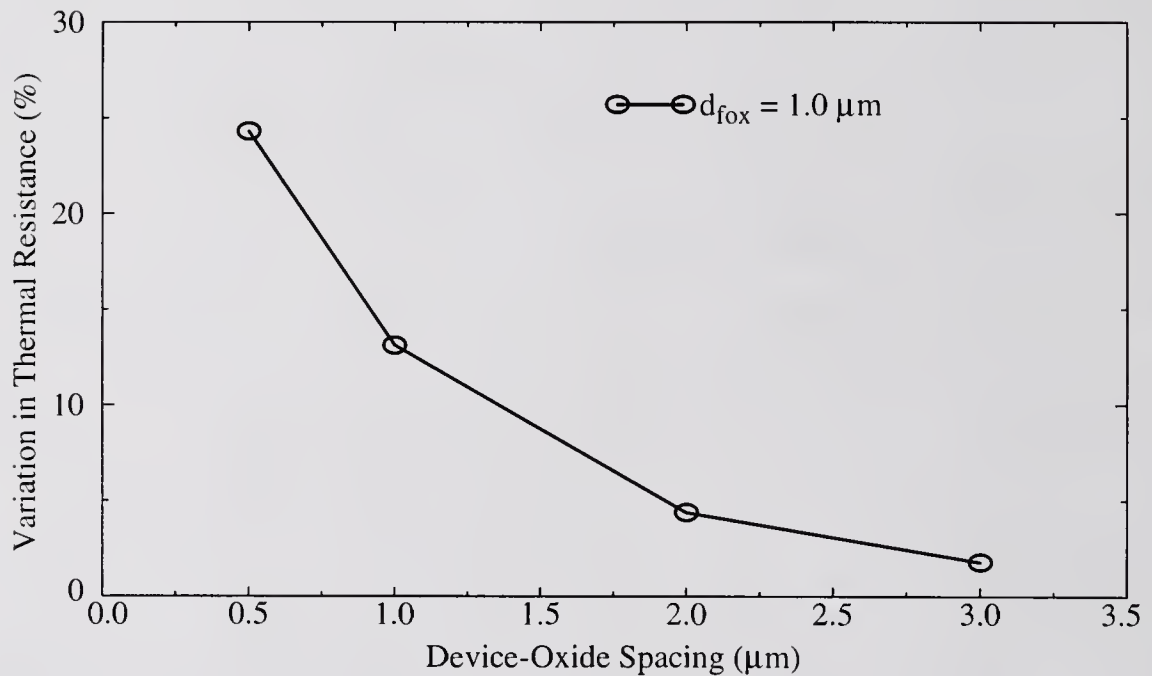
(b)

Figure 2.9   ANSYS simulations showing the effect of recessed LOCOS isolation on the thermal resistance. The device specifications are $L = 2\ \mu m$, $W = 1\ \mu m$, $D = 0.35\ \mu m$ and $H = 0.35\ \mu m$: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for the isolation and the thermal resistance assuming a homogeneous substrate, plotted as a function of the edge-to-edge spacing between the emitter and the isolation.

had a thickness, d_{fox}, of one micron. Figure 2.9b compares the FE simulations with and without the isolation. The effect of the LOCOS can be significant at small device-isolation spacings, but decreases as the isolation is moved away from the active device region. In Figure 2.8a, the portions of the LOCOS structure close to the emitter have the largest effect on the thermal response, since they directly restrict the lateral heat flow away from the device. A number of manufacturers are using thinner standard or semi-recessed LOCOS (0.3 to 0.6 µm) combined with junction isolation to reduce fabrication times and improve compatibility with existing MOS technologies [Klo93], [Pru94]. The thinner oxides have a smaller effect on the thermal resistance, and therefore, the FE simulations can be considered worst-case.

U-groove isolation differs from LOCOS in that trenches are etched directly into the substrate and then back-filled with oxide and polysilicon. The depth of the trench, d_{tr}, is typically on the order of 3 µm [Yam93, Ona95], but has been as large as 5 µm [Del91]; the width of the trench is generally in a range from 0.6 to 1.5 µm [Del91, Yam93, Ona95, Shi96]. U-groove trenches will typically have a surface LOCOS layer, but the thickness of this layer is usually no greater than 0.1 to 0.15 µm [Yam93], since the isolation is mainly achieved by the trench. Figure 2.10a shows the FE model for devices with U-groove isolation. The thickness of each fill layer (d_{trox} for oxide and d_{poly} for polysilicon) in the U-groove was assumed to be uniform. Figure 2.10b compares the FE simulations with and without the isolation. The effect of the U-groove isolation on the thermal resistance is greater for small device-isolation spacings than in LOCOS due to the larger depth of the trench.
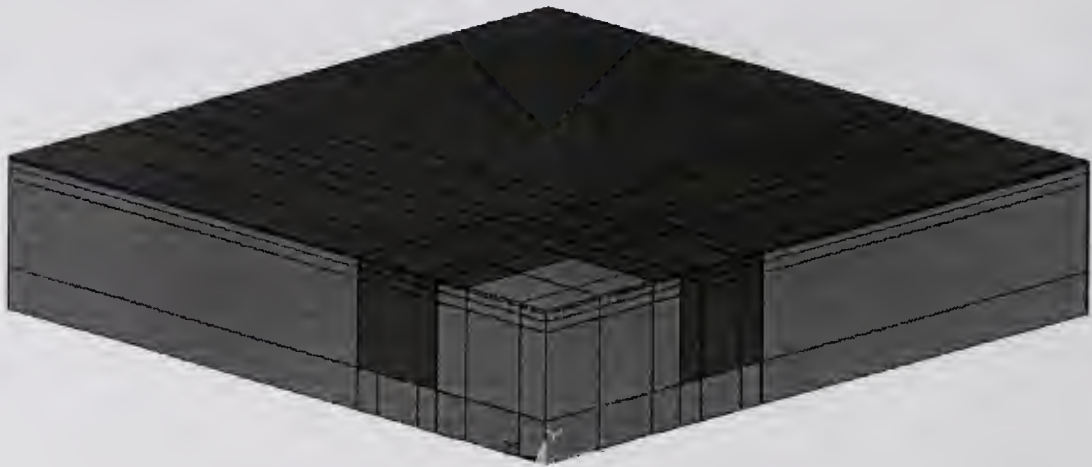
(a)

(b)
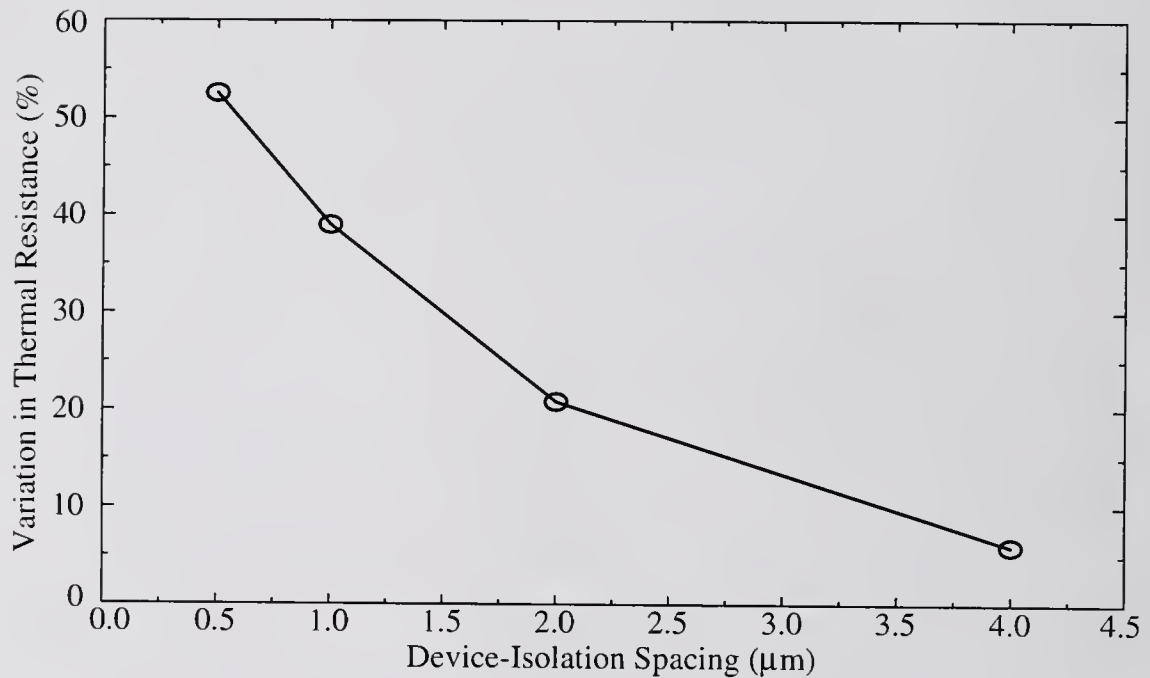
Figure 2.10   ANSYS simulations showing the effect of U-groove isolation on the thermal resistance. The device specifications are L = 2 μm, W = 1 μm,   D = 0.35 μm   and   H = 0.35 μm.   The   U-groove specifications are d_{tr} = 3.5 μm, d_{trox} = 0.38 μm and d_{poly} = 0.75 μm: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for the isolation and the thermal resistance assuming a homogeneous substrate, plotted as a function of the edge-to-edge spacing between the emitter and the U-groove isolation.

However, these results represent the worst case, since in an actual device the active region is not immediately flanked by the U-groove on all sides.

Transient thermal simulations of the FE models shown in Figure 2.9a and Figure 2.10a were used to examine the effects of both LOCOS and U-groove isolation on the transient thermal response. ANSYS was used to simulate the device with and without the isolation structures; the results are shown in Figure 2.11. As the heat travels laterally and reaches the edges of the isolation structure, the response accounting for the isolation begins to deviate from the response without the isolation. The time for the heat to reach the edges of the isolation structures can be approximated as the square of the device-oxide separation (1 μm) divided by the thermal diffusivity of the substrate material; the resulting time is on the order of ten nanoseconds, which agrees with the simulations of both the LOCOS and U-groove isolation. The oxide used in the isolation structures restricts the lateral flow of heat away from the device and in both cases results in a larger temperature rise.

Based on the results of the 3-D FE simulations, advanced isolation structures such as recessed LOCOS and U-groove can considerably increase the thermal impedance of bipolar devices. In most cases, the bulk bipolar model will tend to under-predict both the steady-state and transient thermal response of devices fabricated with oxide-based isolation structures. The error in the model will be the greatest for advanced, highly-scaled devices fabricated with deep trench isolation.
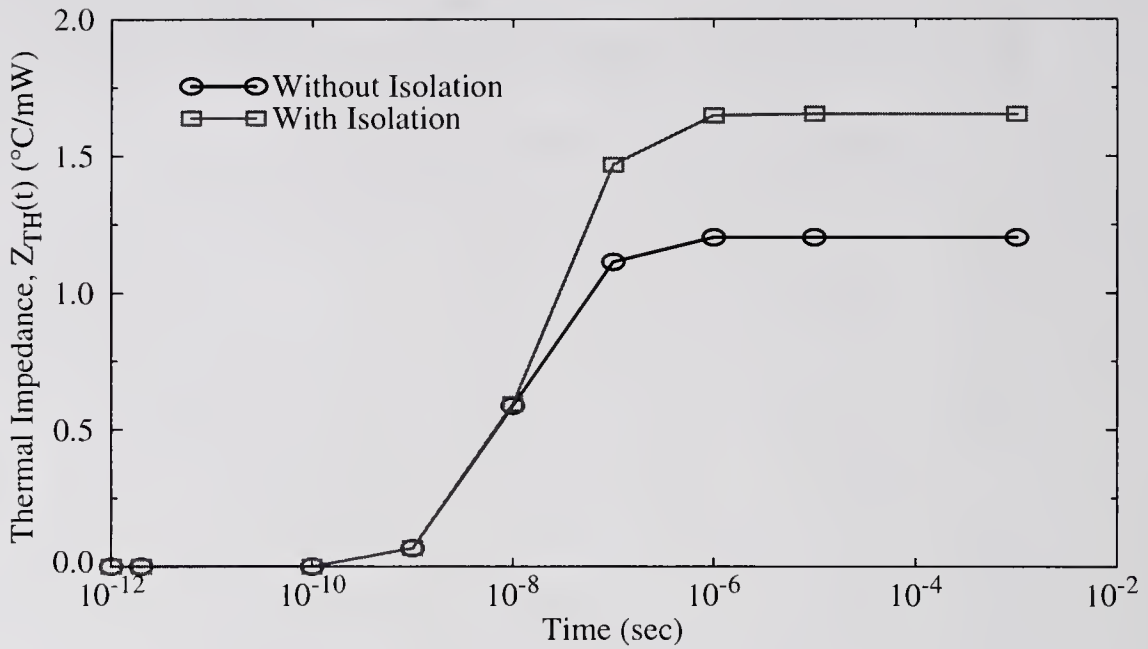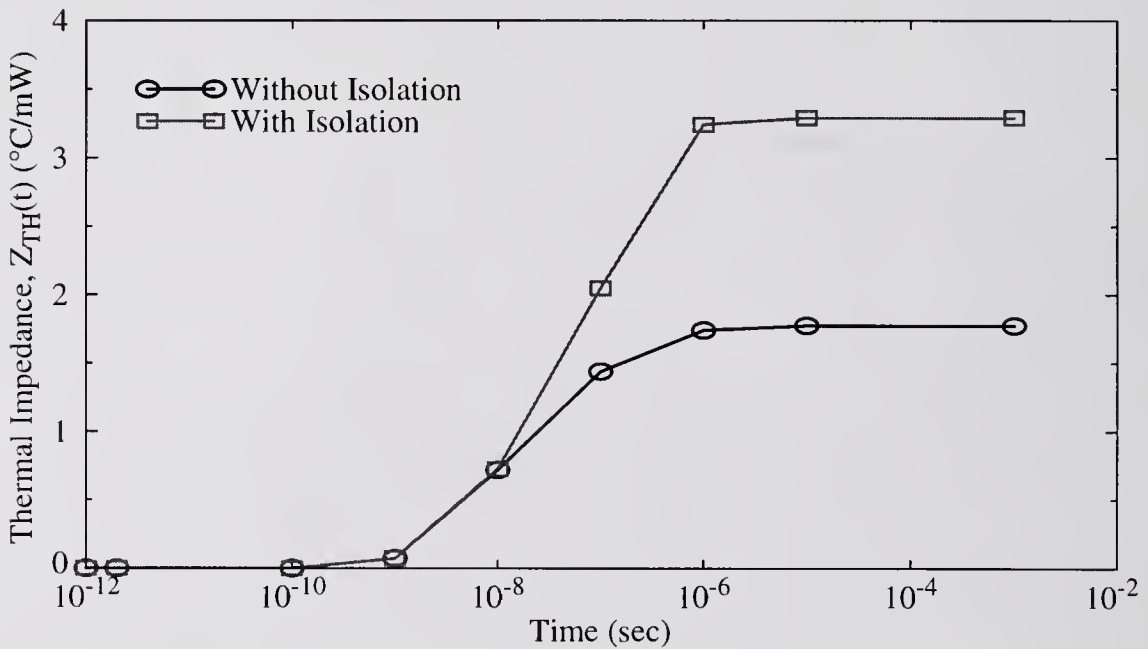
Figure 2.11   ANSYS simulations showing the effect of advanced isolation
structures on the transient thermal impedance. The device-isolation
spacing is one micron. The specifications for the device structure are
$L = 2$ μm, $W = 1$ μm, $D = 0.35$ μm and $H = 0.35$ μm: a) Recessed
LOCOS or BOX isolation with $d_{fox} = 1.0$ μm; b) U-groove isolation
with $d_{tr} = 3.5$ μm, $d_{trox} = 0.38$ μm and $d_{poly} = 0.75$ μm.

2.3 Verification of the Single-Emitter Thermal Impedance Model

To verify the thermal impedance model, three-dimensional (3-D) finite-element (FE) simulations of a junction-isolated BJT were performed using ANSYS. Interconnect metallization was neglected and the substrate was assumed to be homogeneous silicon with the bulk properties given in Table 2.1. The FE simulations were evaluated at the surface corner of the emitter and compared to the single-emitter thermal impedance model given by equation (2.21); the results are shown in Figure 2.12. The analytic model agrees closely, for both the steady-state and the transient, with the 3-D FE simulations for both device geometries. The predicted values for the steady-state thermal resistance agree within twelve percent of the FE simulations. The error can be partially attributed to numerical error associated with the FE mesh.

The model was also compared to measured thermal impedances. The measured data were extracted using the base-thermometry technique developed by Zweidinger et al. [Zwe96]. Figure 2.13, Figure 2.14 and Figure 2.15 compare the measured and simulated data for the transient thermal impedance of Harris HBC bulk BJTs. The thermal impedance model does a good job of predicting the steady-state thermal resistance, with no more than a 20% error between the model and the measurements. The model given by (2.13) tends to overpredict the transient response. As shown with the ANSYS simulations in Figure 2.7, this discrepancy can be attributed to the model's neglect of the emitter metallization. When the thermal impedance of the emitter interconnect is accounted for, with $\delta_{met} = 50$ μm extracted
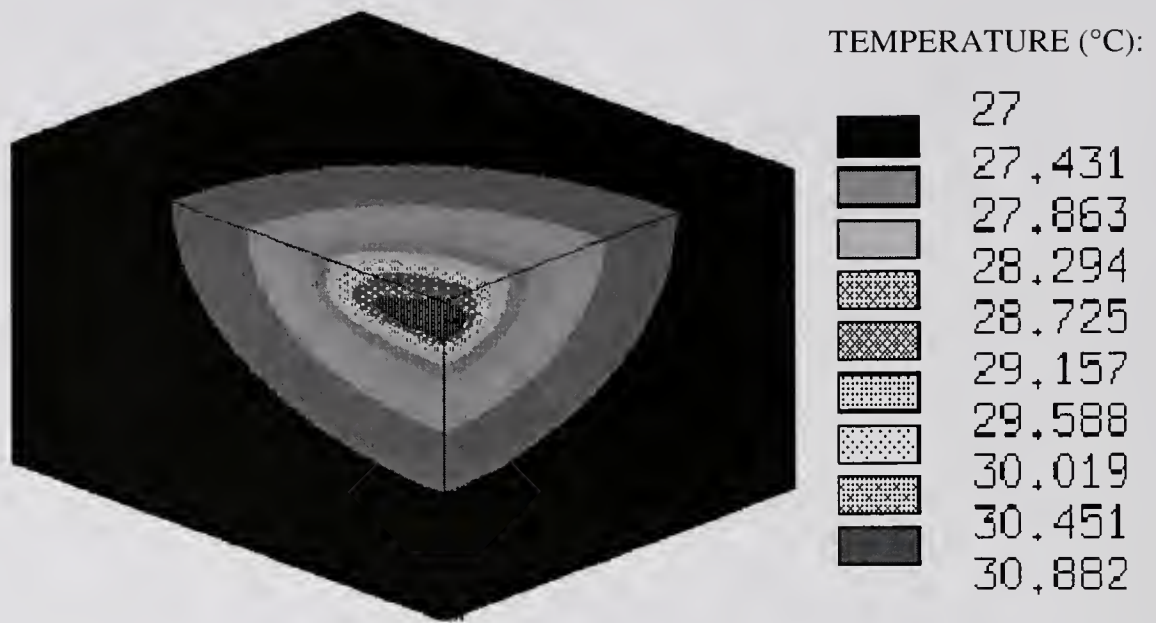
(a)
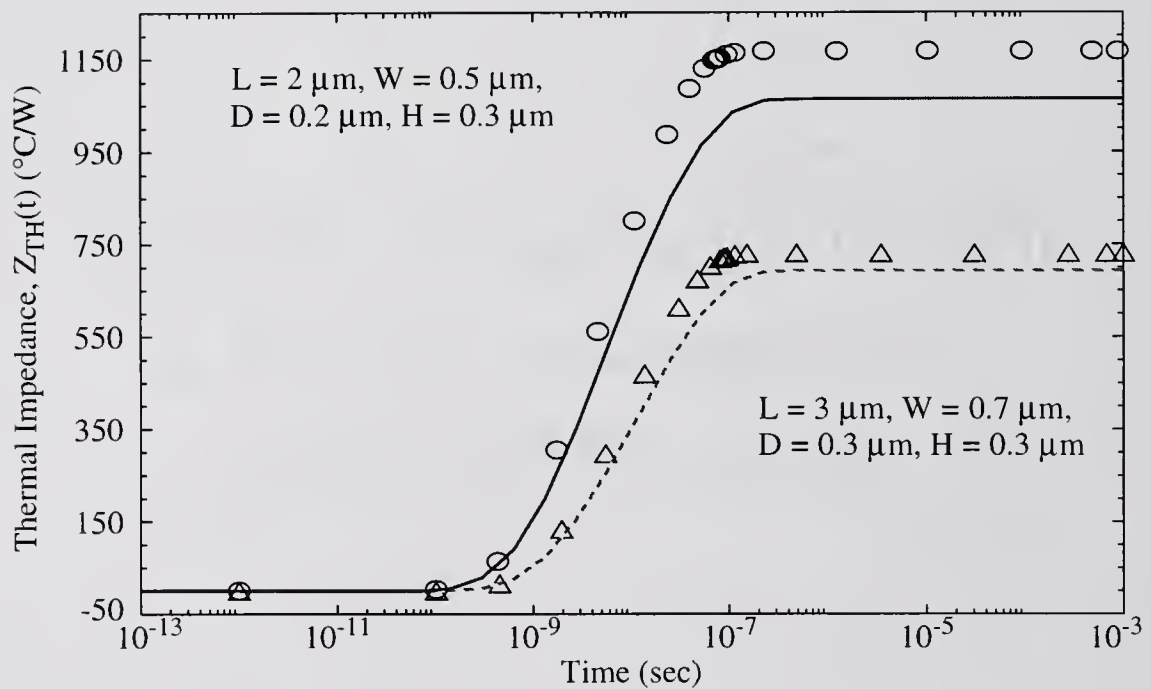
(b)

Figure 2.12   The transient thermal impedance simulated with ANSYS and
calculated with the bulk, single-emitter model: a) The finite-element
model simulated with ANSYS for P = 1.8 mW; b) comparison of
thermal impedances simulated with ANSYS (symbols) and the thermal
impedance model (lines).

(a)

(b)
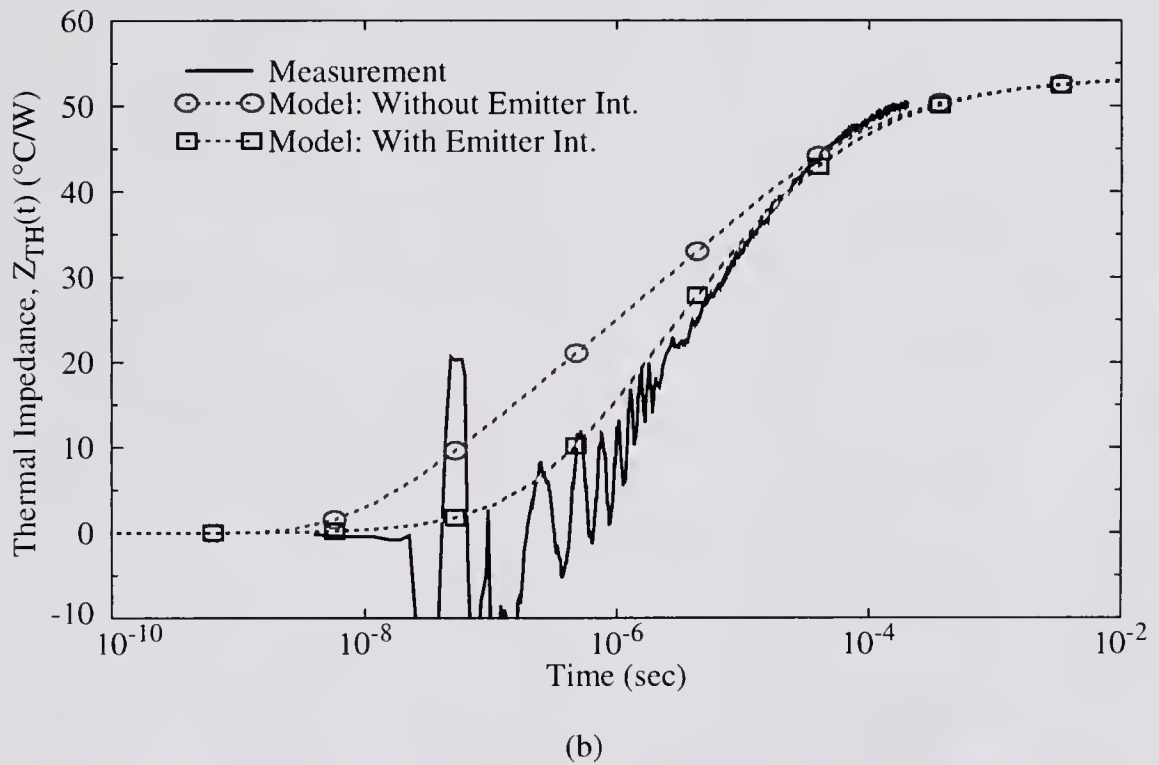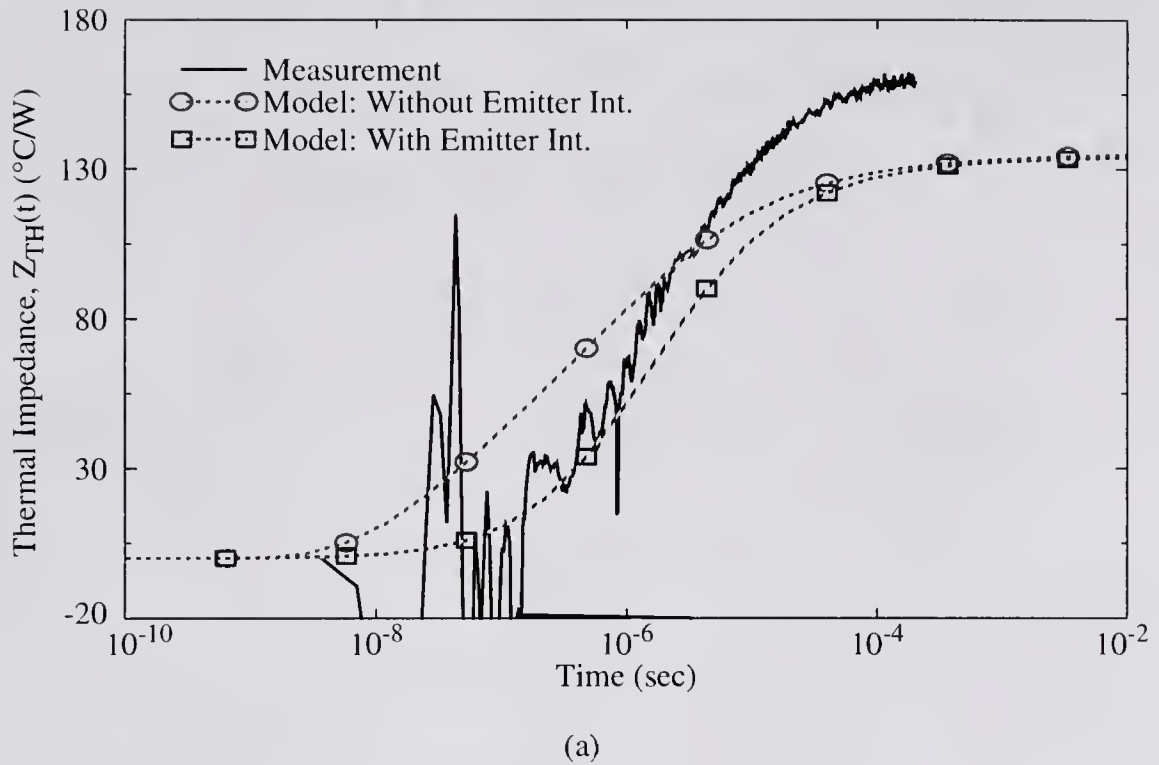
Figure 2.13   Measured and simulated data for the transient thermal impedance of Harris HBC bulk BJT's with $W = 2\ \mu m$: a) $L = 30\ \mu m$; b) $L = 100\ \mu m$.
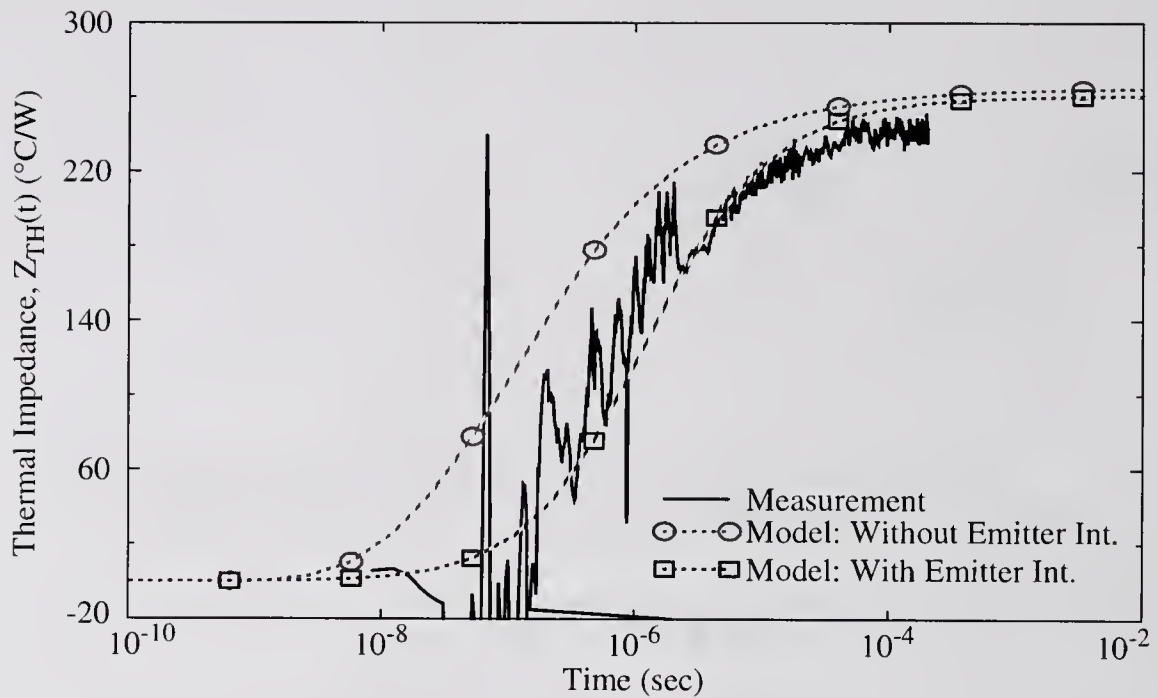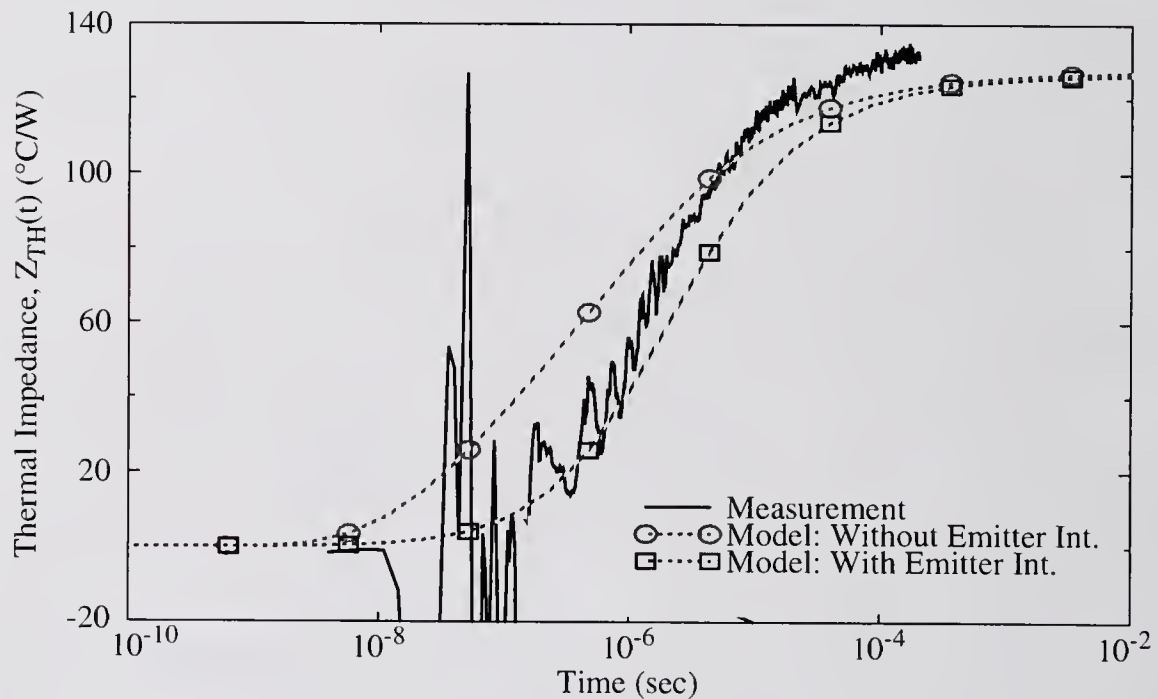
Figure 2.14   Measured and simulated data for the transient thermal impedance of Harris HBC bulk BJT's with W = 3 μm: a) L = 10 μm; b) L = 30 μm.
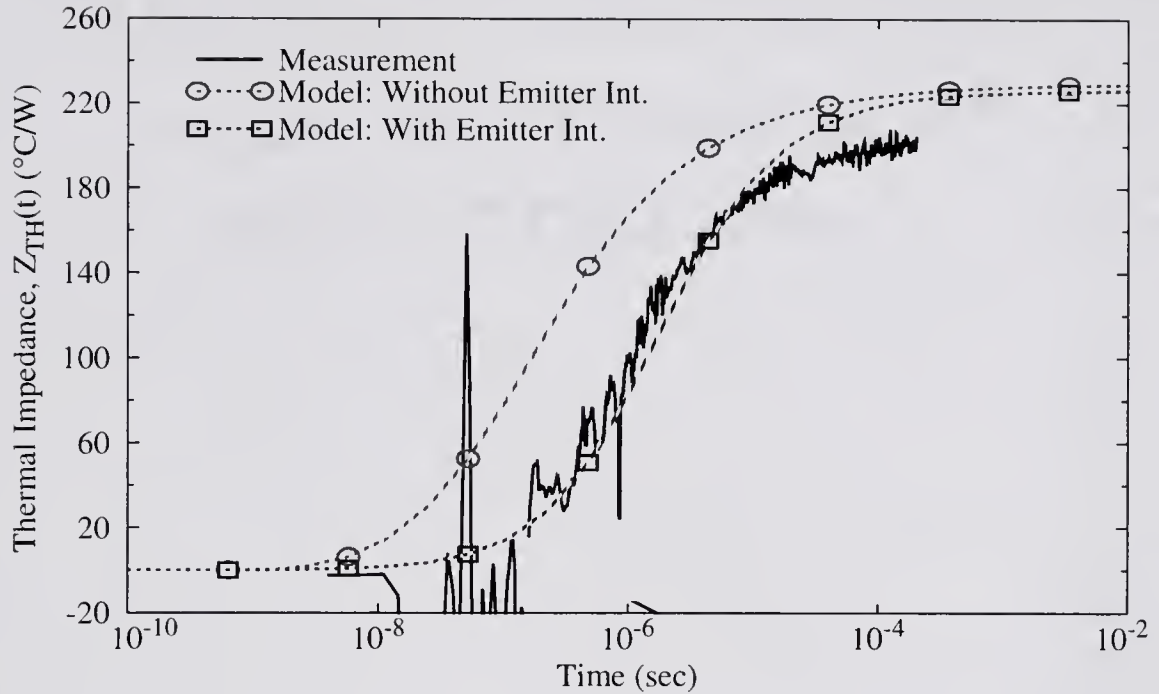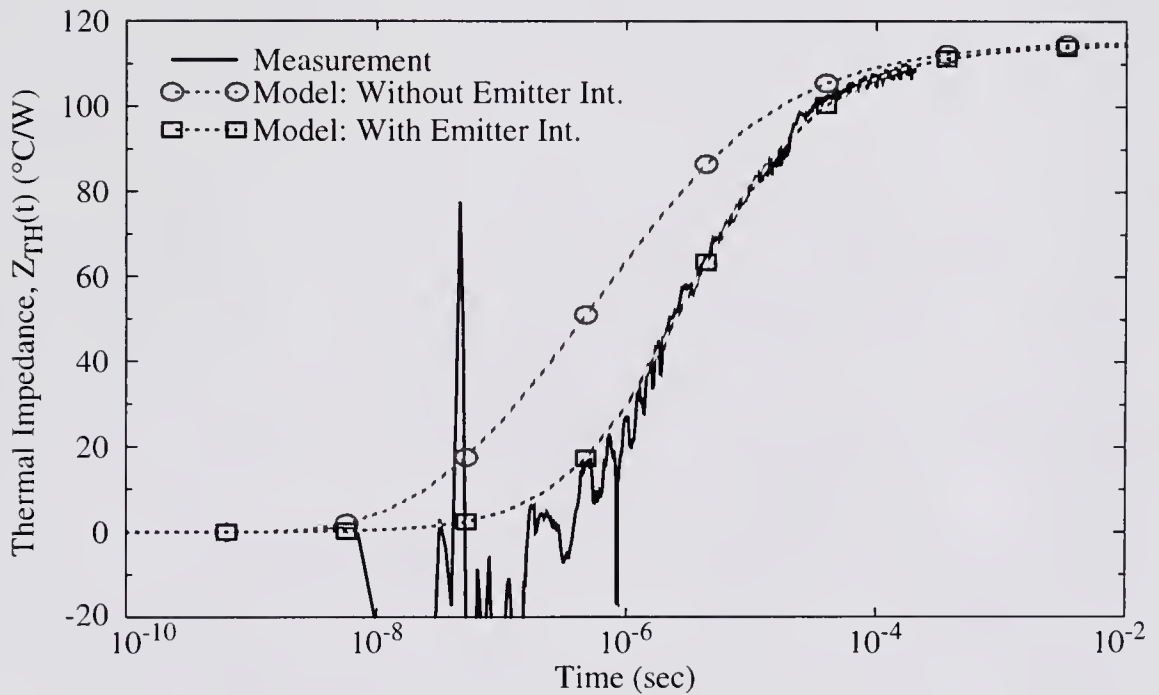
Figure 2.15 Measured and simulated data for the transient thermal impedance of Harris HBC bulk BJT's with W = 5 μm: a) L = 10 μm; b) L = 30 μm.

from the measurements, the model provides a more accurate representation of the transient thermal response.

2.4 Derivation of the Multiple-Emitter BJT/HBT Thermal Impedance Model

The thermal impedance model for bulk MEBJT/MEHBT's is an extension of the single-emitter model. A multiple-emitter device consists of single-emitter devices placed adjacent to each other along their lengths. Since there are multiple devices (referred to as "emitter fingers") that are thermally coupled through the substrate operating in close proximity, the temperature rise in each emitter finger is affected not only by its own power dissipation, but also by the power dissipated by its neighbors. The heat conduction equation, (2.1), is linear, so superposition can be used to calculate the total temperature rise in the device. The equation for the temperature rise can then be manipulated to provide expressions for the effective temperature rise in each individual finger.

Figure 2.16 illustrates the simplified multiple-emitter device geometry assumed for the model derivation. The substrate is represented by a homogeneous semi-infinite half-space with an adiabatic top surface with multiple imbedded heat sources. The emitter fingers are assumed to be uniform in size and shape, with width W and length L. Each finger has a corresponding heat source, due to an assumed uniform power generation in the base/collector SCR; each heat source has a thickness H and is displaced a distance D below the surface of the substrate. As with the single-emitter model, D is assumed to equal the depth of the base/collector junction
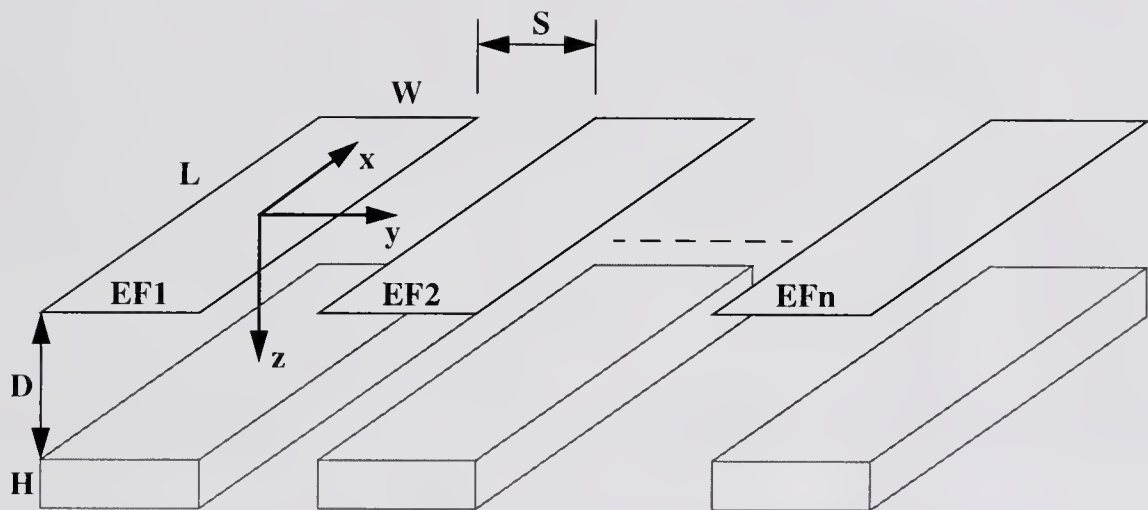
Figure 2.16 The simplified device geometry used to define the solution domain for the bulk, multiple-emitter BJT/HBT thermal impedance model. The substrate is represented by a semi-infinite half-plane with an adiabatic surface. Each emitter finger has a width W and length L and each heat source (the rectangular volumes) is displaced a distance D below the surface of the device and has a thickness H. The distance D is equivalent to the depth of the base/collector junction and the thickness H is approximated by the thickness of the base/collector SCR. The emitter fingers are uniformly spaced with and edge-to-edge separation distance S.

and H can be calculated using equations (2.14) and (2.15). The edge-to-edge separation, S, between adjacent fingers is assumed to be uniform.

The Green's function technique can be employed to find the temperature rise within the device. By applying superposition, the solution is expressed as the sum of the Green's function solutions for the multiple heat sources

$$\Delta T(x, y, z, t) = \sum_{1}^{n} \frac{\alpha}{k} \int_{t'=0}^{t} dt' \int_{V} G(x, y, z, t | x', y', z', t') g(x', y', z', t') dv'. \qquad (2.29)$$

where $G(x, y, z, t | x', y', z', t')$ is given by equation (2.8). The summation accounts for the integration over the different spatial coordinates of each heat source. To clarify the derivation, certain conventions and definitions can be established. The origin for the coordinate system is fixed at the center of the left-most finger at the surface of the device. A device is considered to have a total of n emitter fingers and a reference order is established with the fingers numbered sequentially starting from the left-most finger. The character j, where $j = 1 \rightarrow n$, is used to reference a specific emitter finger. The i-th neighbor (where $i = 1 \rightarrow n-1$) of a given emitter finger, EFj, is defined as a finger situated an edge-to-edge distance $[iS + (i-1)W]$ away on either side. Using equations (2.8) and (2.29), the temperature rise at any point in the device- -assuming a step increase in power at $t' = 0$ for each finger--is given by

$$\Delta T(x, y, z, t) = \sum_{j=1}^{n} \int_t \frac{P_j dt}{8\rho c V} \left[\mathrm{erf}\left(\frac{L/2 + x}{\sqrt{4\alpha t}}\right) + \mathrm{erf}\left(\frac{L/2 - x}{\sqrt{4\alpha t}}\right) \right]$$

$$\cdot \left[\mathrm{erf}\left(\frac{y - (2j-3)W/2 - (j-1)S}{\sqrt{4\alpha t}}\right) \right.$$

$$\left. + \mathrm{erf}\left(\frac{-y + (2j-1)W/2 + (j-1)S}{\sqrt{4\alpha t}}\right) \right]$$

$$\cdot \left[\mathrm{erf}\left(\frac{z + D + H}{\sqrt{4\alpha t}}\right) + \mathrm{erf}\left(\frac{-D - z}{\sqrt{4\alpha t}}\right) \right.$$

$$\left. + \mathrm{erf}\left(\frac{z - D}{\sqrt{4\alpha t}}\right) + \mathrm{erf}\left(\frac{D + H - z}{\sqrt{4\alpha t}}\right) \right] \tag{2.30}$$

which accounts for n heat sources, one for each emitter finger EFj.

Equation (2.30) can be manipulated to provide the temperature rise in each emitter finger. As with the single-emitter model, the temperature rise in each finger is represented by a single effective value. To simplify the derivation, symmetry is assumed such that the distance from the effective-temperature point of finger EFj to the heat source of its i-th neighbor, is the same as the distance from the effective-temperature point of the i-th neighbor to the heat source of EFj. This symmetry is attained only for the coordinates $(x, y = [j - 1][S + L], z)$. When (2.30) is evaluated at each of these points, the model is reciprocal and the effective temperature rise in each emitter finger, ΔT_{EFj}, can be expressed as

$$
\begin{bmatrix} \Delta T_{EF1} \\ \Delta T_{EF2} \\ \cdot \\ \cdot \\ \cdot \\ \Delta T_{EFn} \end{bmatrix} = \begin{bmatrix} Z_S & Z_{C1} & \cdots & Z_{C(n-1)} \\ Z_{C1} & Z_s & \cdots & Z_{C(n-2)} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ Z_{C(n-1)} & Z_{C(n-2)} & \cdots & Z_S \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \\ \cdot \\ \cdot \\ \cdot \\ P_n \end{bmatrix}
$$

(2.31)

where Z_S is referred to as a self impedance and Z_{Ci} is the i-th coupling impedance.

The system of equations given by (2.31) shows that the temperature rise in each finger is determined by the power dissipation in its own heat source and by the power dissipated in the $(n-1)$ neighboring heat sources. The self impedance is given by

$$
\begin{aligned}
Z_S(x, z, t) = \sum_{j=1}^{n} \int_t \frac{dt}{4\rho c V} & \left[\mathrm{erf}\left(\frac{L/2 + x}{\sqrt{4\alpha t}}\right) + \mathrm{erf}\left(\frac{L/2 - x}{\sqrt{4\alpha t}}\right) \right] \mathrm{erf}\left(\frac{W/2}{\sqrt{4\alpha t}}\right) \\
& \cdot \left[\mathrm{erf}\left(\frac{z + D + H}{\sqrt{4\alpha t}}\right) + \mathrm{erf}\left(\frac{-D - z}{\sqrt{4\alpha t}}\right) \right. \\
& \left. + \mathrm{erf}\left(\frac{z - D}{\sqrt{4\alpha t}}\right) + \mathrm{erf}\left(\frac{D + H - z}{\sqrt{4\alpha t}}\right) \right]
\end{aligned}
$$

(2.32)

and accounts for the portion of the temperature rise in a finger due to that finger's own power dissipation. The i-th coupling impedance can be expressed as

$$Z_{Ci}(x, z, t) = \sum_{j=1}^{n} \int_t \frac{dt}{8\rho c V} \left[erf\left(\frac{L/2 + x}{\sqrt{4\alpha t}}\right) + erf\left(\frac{L/2 - x}{\sqrt{4\alpha t}}\right) \right]$$

$$\cdot \left[erf\left(\frac{(W/2 + i(W + S))}{\sqrt{4\alpha t}}\right) + erf\left(\frac{(W/2 - i(W + S))}{\sqrt{4\alpha t}}\right) \right]$$

$$\cdot \left[erf\left(\frac{z + D + H}{\sqrt{4\alpha t}}\right) + erf\left(\frac{-D - z}{\sqrt{4\alpha t}}\right) \right.$$

$$\left. + erf\left(\frac{z - D}{\sqrt{4\alpha t}}\right) + erf\left(\frac{D + H - z}{\sqrt{4\alpha t}}\right) \right] \tag{2.33}$$

and accounts for the portion of the temperature rise in a finger due to the power dissipated by its i-th neighbor. Thus, for the assumed symmetry, the thermal impedance of a device with n emitter fingers can be described with a single self impedance and $(n - 1)$ coupling impedances.

Equations (2.32) and (2.33) should be evaluated at a single point to represent the temperature rise of each finger by a single effective value. To keep the MEBJT/MEHBT model as similar as possible to the single-emitter model, the points $x = L/2$ and $z = 0$ are used, giving

$$Z_S(t) = \int_t \frac{1}{2\rho c V} erf\left(\frac{L}{\sqrt{4\alpha t}}\right) erf\left(\frac{W/2}{\sqrt{4\alpha t}}\right) \left[erf\left(\frac{D + H}{\sqrt{4\alpha t}}\right) - erf\left(\frac{D}{\sqrt{4\alpha t}}\right) \right] dt \tag{2.34}$$

and

$$Z_{Ci}(t) = \int_t \frac{1}{4\rho c V} \left[erf\left(\frac{L}{\sqrt{4\alpha t}}\right) \right]$$

$$\cdot \left[erf\left(\frac{(W/2 + i(W + S))}{\sqrt{4\alpha t}}\right) + erf\left(\frac{(W/2 - i(W + S))}{\sqrt{4\alpha t}}\right) \right]$$

$$\cdot \left[erf\left(\frac{D + H}{\sqrt{4\alpha t}}\right) - erf\left(\frac{D}{\sqrt{4\alpha t}}\right) \right] dt \tag{2.35}$$

as the final expressions for the self and coupling impedances. Therefore, the overall self-heating on the scale of the entire device can be described by the self-heating and thermal interactions on the smaller scale of each individual emitter finger.

Accounting for a finite wafer thickness can be important for multiple-emitter devices since the coupling impedances decrease as the wafer thickness is reduced [Daw94], [Hat95]. The multiple-emitter model can be modified in a similar fashion as that for the single-emitter model; by simply using equations (2.17) and (2.20) in place of the $G_z(z, t | z', t')$ in equation (2.29), the expressions for Z_S and Z_{Ci} will now account for a finite wafer thickness. Since the multiple-emitter thermal impedance model is simply an extension of the single-emitter model, the effects of interconnect metallization and advanced isolation technologies are not taken into account. Neglecting these portions of the overall device structure is assumed to affect the multiple-emitter model in the same manner, and to the same extent, as to which it affects the single-emitter model.

2.5 Verification of the Multiple-Emitter Thermal Impedance Model

Two-dimensional (2-D) finite-element (FE) simulations of a junction-isolated, three-finger BJT were performed using ANSYS to verify the multiple-emitter thermal impedance model. Two-dimensional FE simulations were used instead of 3-D simulations due to limitations of the available version of ANSYS. The validity of comparing 2-D FE simulations to a 3-D analytic model is established by evaluating the thermal impedance model derived for both two and three dimensions.

Figure 2.17a compares the results, which show that the 3-D the 2-D models converge for long devices. The difference in the predicted thermal resistance values decreases from 22% to less than 1% as the length of the device is increased from 50 μm to 800 μm. Consequently, the 2-D FE simulations can verify the 3-D thermal impedance model evaluated for devices with long emitters. The 2-D FE simulations do not verify the model for shorter devices where the heat flow becomes three-dimensional. However, the verification of the single-emitter thermal impedance model for 3-D heat flow can be assumed to also verify the MEBJT model. This assumption is reasonable since the physics that describe the single-emitter model also apply to the MEBJT thermal impedance model.

The physical device was assumed to be symmetric so that the FE model represented only half of the device. Figure 2.17b shows an illustration of the FE model simulated with ANSYS. Interconnect metallization was neglected and the substrate was assumed to be homogeneous silicon with the bulk properties given in Table 2.1. The bottom and exterior side of the substrate were held at a constant ambient temperature while the top surface and the interior side were assumed to be adiabatic. The FE simulations were compared to the multiple-emitter thermal impedance model given by equations (2.34) and (2.35), accounting for a finite substrate thickness; the results are shown in Figure 2.18. The analytic model agrees well, in both the steady-state and transient, with the 3-D FE simulations of the self and coupling impedances. The predicted values for the steady-state thermal resistance agree within three percent of the FE results, which is within the expected error of the model and the numerical simulations.

(a)

(b)
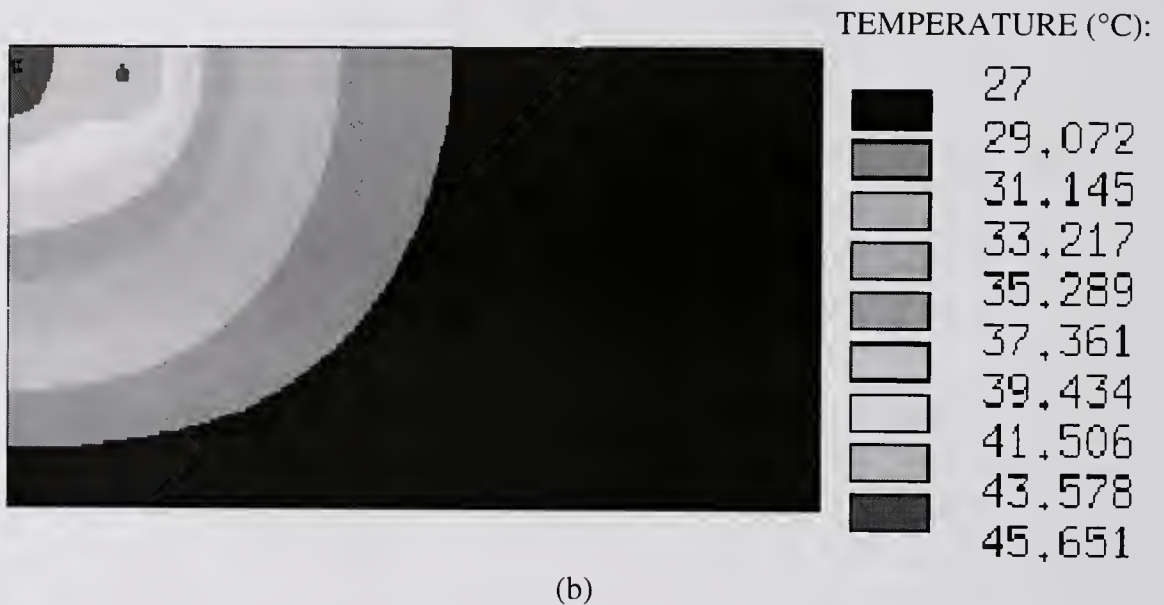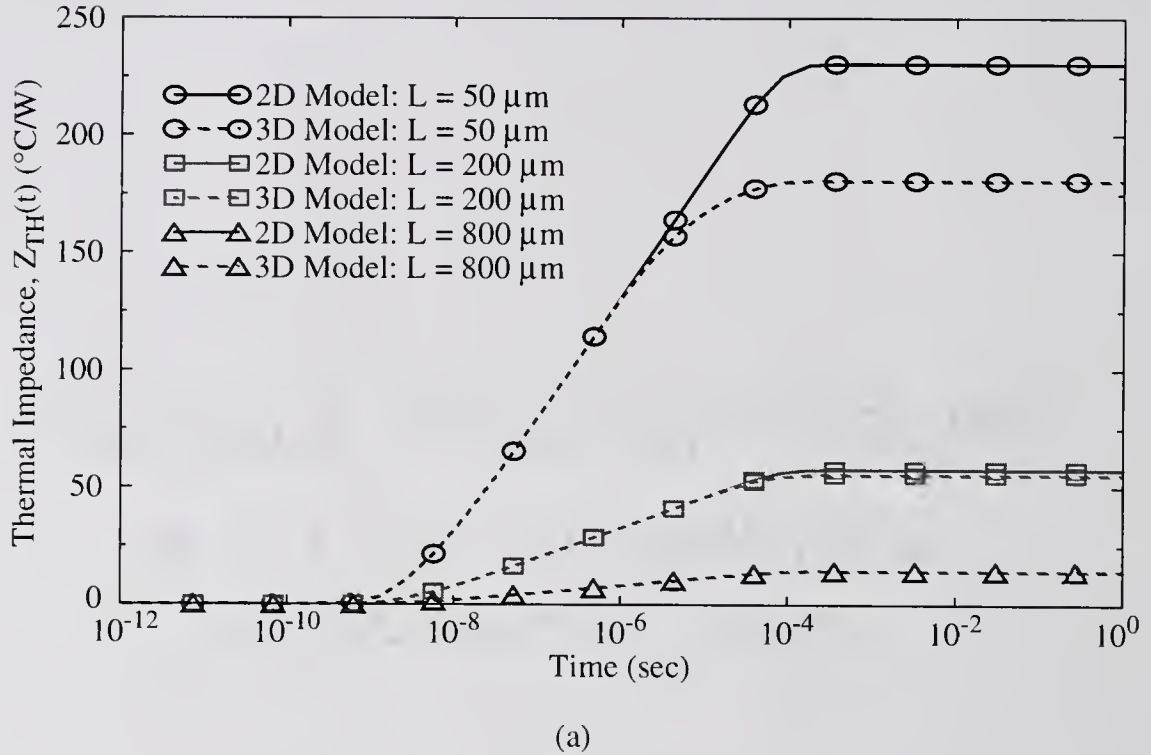
Figure 2.17   Two-dimensional heat flow in multiple-emitter bipolar transistors: a)
A comparison of the three-dimensional thermal impedance model to a
two-dimensional model for W = 1 μm, D = 0.5 μm and H = 0.5 μm;
b) the finite-element model simulated with ANSYS for P = 500 μW.
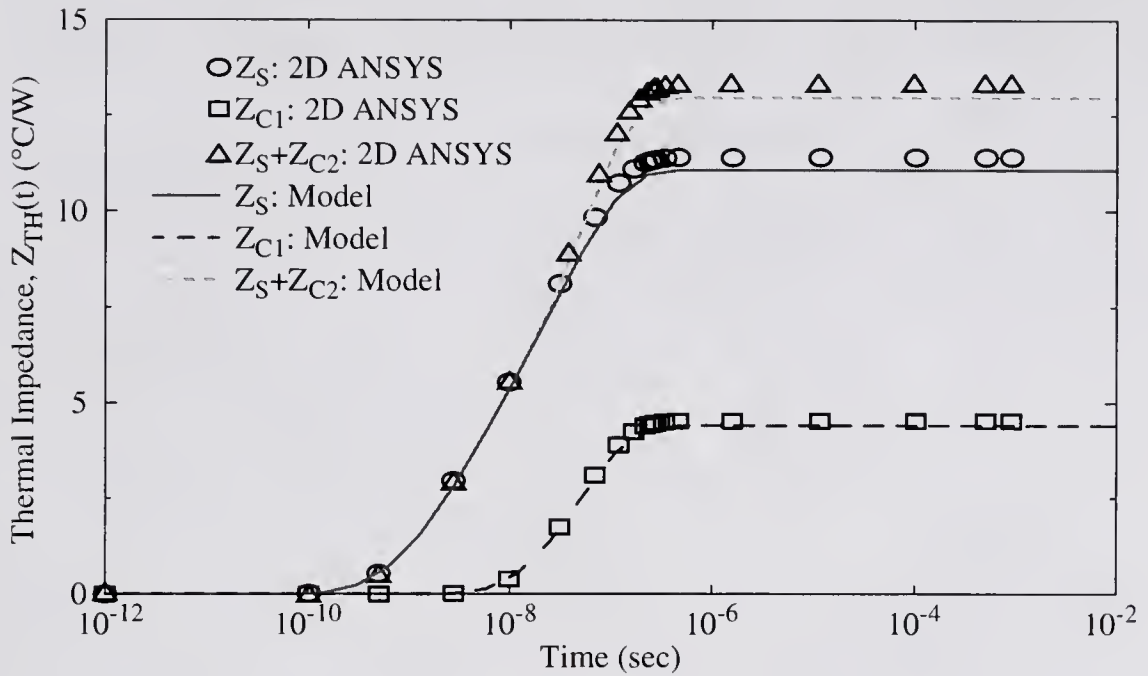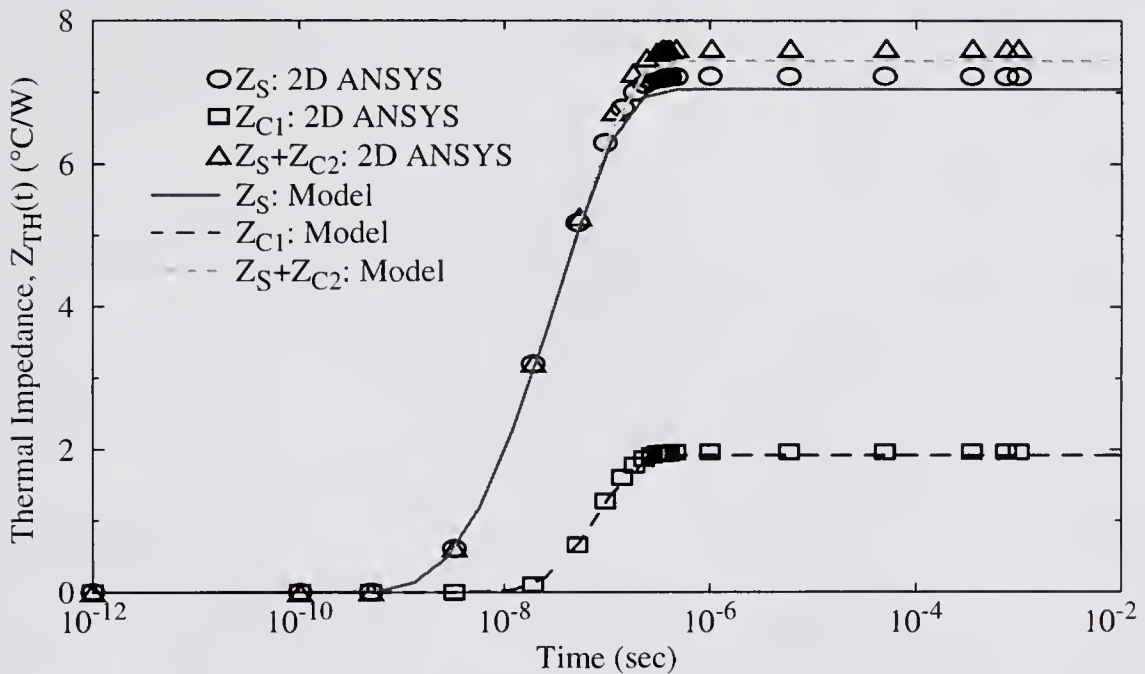
Figure 2.18   The transient thermal impedance simulated with ANSYS and calculated with the bulk, multiple-emitter model for L = 500 μm and S = 1 μm: a) W = 1 μm, D = 0.2 μm, H = 0.3 μm; b) W = 3 μm, D = 0.5 μm, H = 0.5 μm.

2.6 Summary

A thermal impedance model for bulk single-emitter BJT/HBT's was presented and then extended for devices with multiple emitter fingers. The model was shown to agree reasonably well with three-dimensional finite-element simulations and measurements of junction-isolated devices. The effects of interconnect metallization and advanced isolation technologies on the thermal impedance were investigated; a simple model for the thermal impedance of the emitter interconnect was demonstrated. The results suggest that the model can be expected to provide reasonable predictions for the thermal impedance of junction-isolated devices. However, for highly-scaled devices, the effects of advanced isolation can be significant and the accuracy of the model will decline. Methods for modeling the effects of isolation structures are proposed in Chapter Eight.

CHAPTER 3
A CIRCUIT MODEL FOR THERMAL COUPLING AND A LUMPED ELECTROTHERMAL MODEL FOR BULK MULTIPLE-EMITTER BIPOLAR TRANSISTORS

3.1 Introduction

Due to increased interest in the role of thermal effects in device and circuit operation, especially for silicon-on-insulator (SOI) and heterojunction technologies, circuit simulators and compact device models have been modified to account for the dynamic temperature response within a device [McA92, Fox93b, Fos95]. Most of the implementations have been applied to the case of self-heating, where a device's effective operating temperature (EOT) depends on its power dissipation only. In many circuits and some devices, such as multiple-emitter bipolar transistors, a number of devices can operate in close proximity. Under such conditions, the EOT of a device is no longer determined solely by its own power dissipation but also depends on the operation of its neighbors. Therefore, not only must circuit simulators (and compact device models) be able to model dynamic self-heating, they must also be able to model the dynamic thermal coupling between individual devices or portions of one device.

An approach for modeling cross-chip thermal coupling using a circuit simulator was described by Fukahori and Gray [Fuk76]. The thermal coupling between devices in an arbitrary circuit was modeled using a finite difference

77

technique. The semiconductor substrate was represented by a three-dimensional numerical mesh with equivalent thermal resistances and capacitances. The electrical elements in the circuit (transistors, etc.) were represented by their standard compact circuit models. The values for the lumped thermal components were calculated by discretizing the heat conduction equation using a finite difference approximation. In [Mar93], a similar approach was presented and applied to the simulation of multiple-emitter HBT's. In this case, two-port theory was used to generate a finite two-dimensional resistance network that represented steady-state heat conduction in the substrate. For both applications, the resulting circuit admittance matrix contained elements corresponding to the electrical circuit and also the thermal elements. The Newton-Raphson-like iteration scheme of the modified circuit simulators was then used to solve the coupled electrothermal problem. To simulate both inter-device thermal coupling and self-heating using this method, a large number of thermal nodes is required; therefore, this approach can drastically increase simulation time.

Chapter One described a common method for using thermal impedances to efficiently model self-heating in circuit simulators. A logical progression would be to expand this method to account for thermal coupling between devices. Such an approach can provide a more efficient alternative to the semi-numerical methods mentioned above. This technique was applied by Moinian et al. for modeling cross-substrate thermal coupling in bipolar circuits [Moi94], and by Baureis for modeling multiple-emitter HBT's [Bau94]. However, their circuit implementations did not correctly represent the thermal interactions between (or within) the devices. The shortcoming of the coupling model used in these works is discussed in this chapter.

A circuit model is then presented which correctly models thermal coupling and is compatible with the self-heating circuit model described in Chapter One.

Once a valid circuit model for thermal coupling has been developed, it can be used with the multiple-emitter thermal impedance model to perform both steady-state and dynamic electrothermal simulations of multiple-emitter BJT/HBT's. The multiple-emitter thermal impedance model expresses the self-heating of an entire device as the sum of the thermal actions and interactions of the individual emitter fingers. The thermal model structure requires that a single multiple-emitter device be represented by multiple compact device models. While this configuration allows for the examination of the EOT of each finger in a device, for devices with a large number of emitter fingers, the overall electrothermal network can become complex enough to make moderate- to large-scale circuit simulations impractical. To make the multiple-emitter electrothermal model more suitable for circuit simulation, its complexity can be reduced by representing the overall thermal response of the device by a lumped thermal impedance. The lumped thermal impedance is generated by applying the measurement approach developed by Zweidinger et al. to the simulation of the complete electrothermal model [Zwe96]. The thermal impedance extraction technique is briefly reviewed in this chapter. The lumped model generation procedure is then described and the results are compared to the complete electrothermal model.

3.2 A Circuit Model for Thermal Coupling

The EOT of any device in a system of n thermally coupled devices (e.g. a multiple-emitter bipolar transistor with n emitter fingers) can be expressed as

$$T_{DEV1}(t) = [\Delta T_1(t) + \Delta T_{12}(t) + \cdots + \Delta T_{1n}(t)] + T_{amb}$$
$$T_{DEV2}(t) = [\Delta T_{21}(t) + \Delta T_2(t) + \cdots + \Delta T_{2n}(t)] + T_{amb}$$
$$\cdot$$
$$\cdot$$
$$\cdot$$
$$T_{DEVn}(t) = [\Delta T_{n1}(t) + \Delta T_{n2}(t) + \cdots + \Delta T_n(t)] + T_{amb} \tag{3.1}$$

where $\Delta T_j(t) = \pounds^{-1}[Z_{Sj} \cdot P_j(s)]$ and $\Delta T_{ji}(t) = \pounds^{-1}[Z_{Cji} \cdot P_i(s)]$. The impedance Z_{Sj} is the self impedance of device j, and Z_{Cji} represents the coupling impedance between device j and device i. (In general, it is not necessary for Z_{Cji} to equal Z_{Cij}.) The modifications described in Chapter One allow circuit simulators to model self-heating; therefore, the EOT of each device in a simulation is calculated by

$$T_{DEVj}(t) = \Delta T_j(t) + T_{amb} \tag{3.2}$$

and is independent of its neighbors. The obvious way to expand this technique to account for inter-device thermal coupling would be to simply tie together the temperature nodes of individual devices using coupling impedances; this approach was used by Baureis and Moinian et al. [Bau94, Moi94]. Figure 3.1 shows an example of such a thermal coupling network for two devices, where $Z_C = Z_{C12} = Z_{C21}$. Unfortunately, simple analysis of the circuit in Figure 3.1 shows that it does not correctly model the expression in (3.1). For example, analyzing
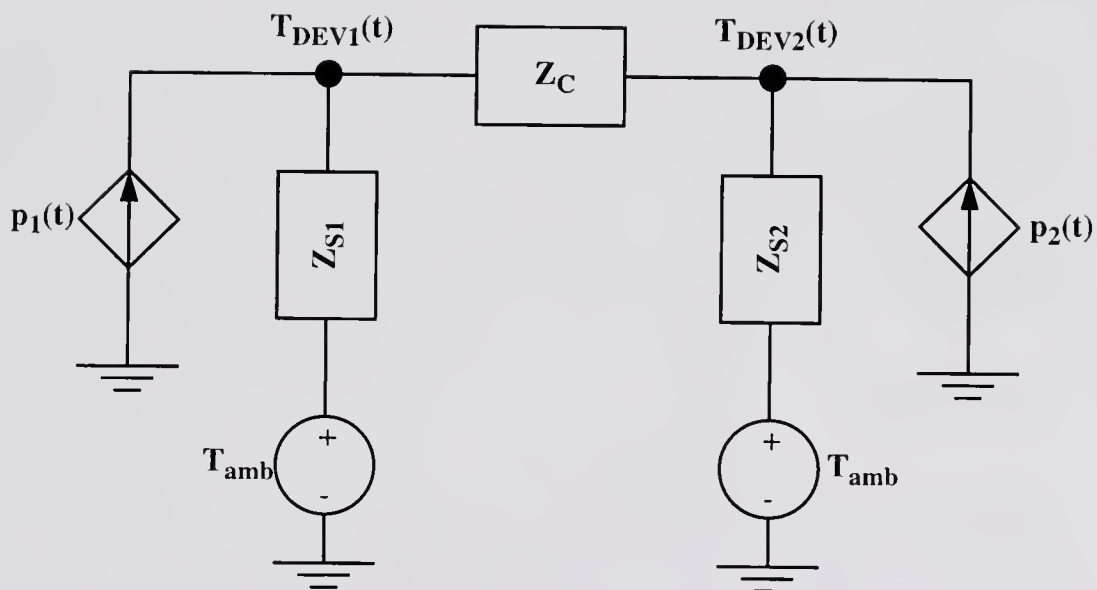
Figure 3.1   A thermal coupling circuit model for two devices. The temperature nodes of the two thermally coupled devices are connected using a thermal coupling impedance.

the circuit in the steady-state limit gives the following expression for the EOT of device 1

$$T_{DEV1} = \frac{R_{S1}(R_C + R_{S2})}{(R_C + R_{S1} + R_{S2})}P_1 + \frac{R_{S1}R_{S2}}{(R_C + R_{S1} + R_{S2})}P_2 + T_{amb}. \qquad (3.3)$$

A similar expression can be derived for the EOT of device 2. The problem with this network formulation is that when individual temperature nodes are connected through an impedance path, the entire network becomes distributed among the coupled devices. The self impedances and coupling impedances, as derived, are not defined to be distributed elements. In a more simplistic view, the network in Figure 3.1 does not properly constrain the paths of the respective device power-currents. The power-current of a given device is divided between its own self impedance and the rest of the network. The portion of that device's power flowing through its neighbor's self impedance has no physical meaning. As a result, the voltages generated at the temperature nodes do not correspond to the correct device temperatures.

To develop a correct circuit representation of (3.1) and avoid the shortcomings of the aforementioned coupling technique, control sources can be utilized in a thermal coupling network group composed of two sub-networks. Each device in a group of n thermally coupled devices has its own network group. Figure 3.2 demonstrates how a thermal coupling network group works. Sub-network A attaches directly to the temperature node of device 1. The current-controlled current source (F_1) in sub-network B has unity gain and is controlled by the current
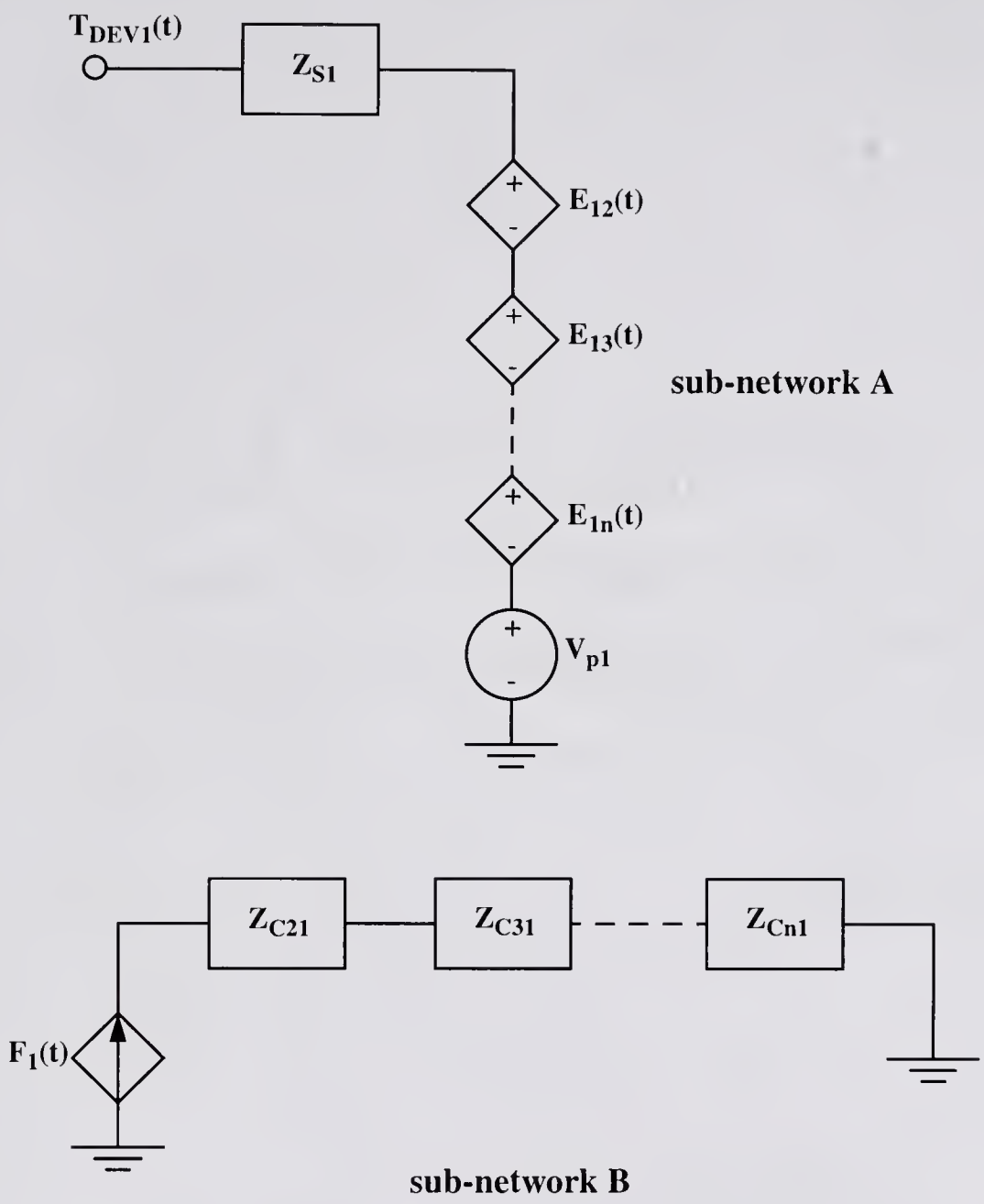
Figure 3.2   A new circuit representation of thermal coupling which is compatible
with the self-heating circuit model. Sub-network A is attached to the
temperature node of a device and represents the total temperature rise
in that device. Sub-network B is used to calculate the thermal coupling
between devices. The voltage-controlled voltage sources (E_{1j})
represent the individual portions of the temperature rise in device 1 due
to the other devices in the circuit. The current-controlled current
source (F_1) models the power dissipation in device 1.

flowing through the voltage source V_{p1} in sub-network A. In this example, V_{p1} is also used to set the reference ambient temperature. The voltage drop across each coupling impedance (Z_{Ci1}) in sub-network B corresponds to the portion of the temperature rise in each device i due to the power dissipation of device 1. The voltage-controlled voltage sources in sub-network A each have unity gain and are used to couple the voltage drops from each sub-network B of the other devices, back to device 1. For example, the value of the voltage source E_{12} is equivalently $E_{12} \equiv \Delta T_{12}(t)$, where Z_{C12} is part of sub-network B of device 2. Therefore, the voltage generated at the terminal of sub-network A corresponds to the EOT of device 1, and is given by the following expression

$$T_{DEV1}(t) = \Delta T_1(t) + \Delta T_{12}(t) + \cdots + \Delta T_{1n}(t) + T_{amb}. \tag{3.4}$$

Similar expressions can be obtained for the EOT's of the other $(n-1)$ devices in the circuit since they each have similar thermal networks.

The thermal coupling model is demonstrated by simulating a five-finger HBT using a version of SPICE 2G.6 modified to model self-heating [Zwe97]. The device characteristics are simulated with and without the thermal coupling between emitter fingers. Figure 3.3 shows the results of the electrothermal simulations. When accounting for the thermal coupling, the current collapse phenomenon commonly observed in HBT's can be simulated [Liu93, Sei93, Liu95b]. Figure 3.3b illustrates how the outer fingers shut down as the middle finger begins to carry all of the current. The importance of modeling the thermal coupling is established by the fact that the collapse phenomenon is not reproduced when the simulations only account for self-
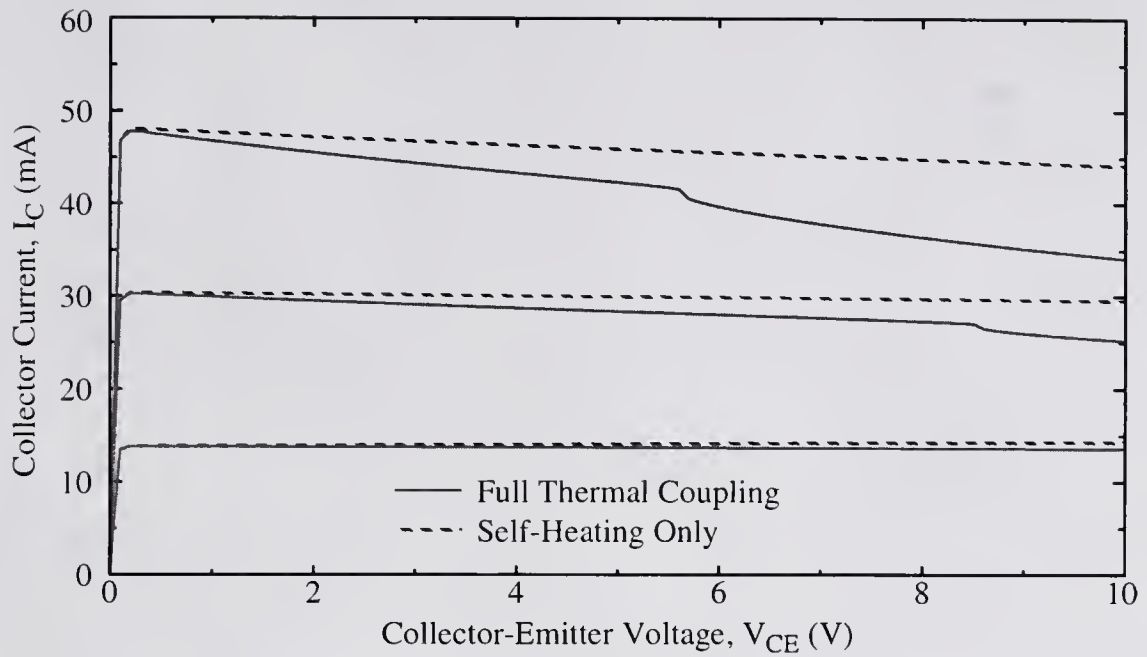
(a)
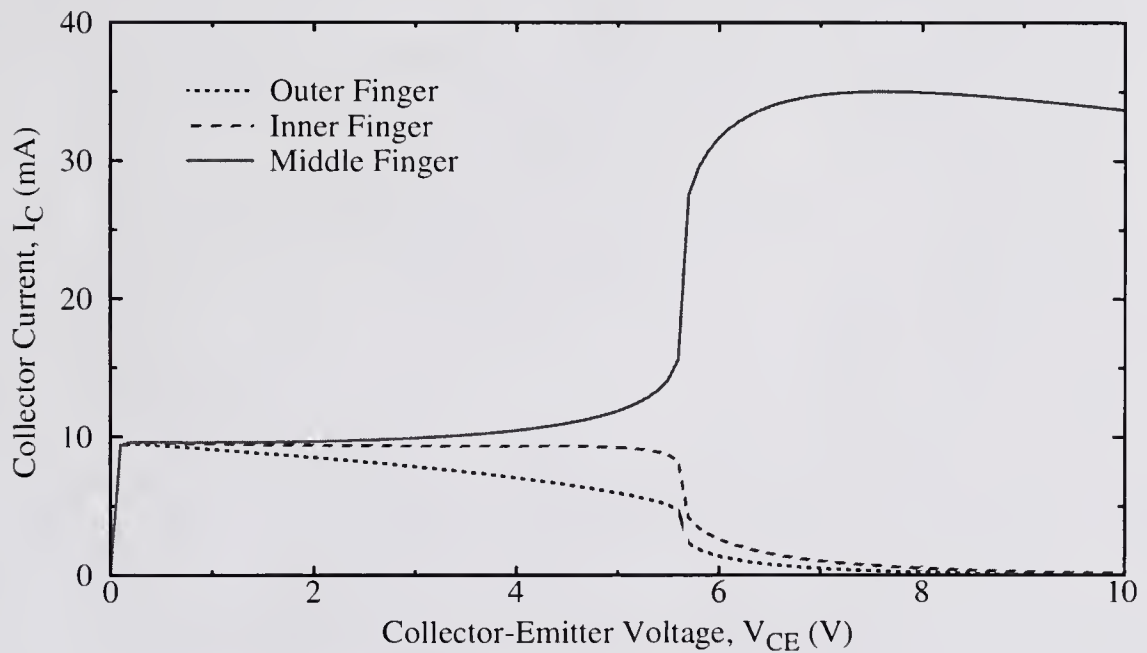
(b)

Figure 3.3   Simulated current characteristics of a five-finger HBT with A_E = 20 x 2 μm^2 for each emitter finger: a) The collector current as a function of collector-emitter voltage for fixed base currents of 1.0, 2.0 and 3.0 mA; b) the collector current distribution in the device with full thermal coupling for I_B = 3.0 mA.

heating. Multiple-emitter devices provide just one application for the thermal coupling model. It can be used on a larger scale for simulating thermal interactions within circuits, and it can be used on a smaller scale. By dividing a single device (or each finger of a device) into multiple sub-cells, the thermal coupling model could be used to simulate the temperature distribution within a device and phenomena such as current constriction [Koe94].

3.3 A Lumped Electrothermal Model for Multiple-Emitter BJT/HBT's

Used together with a compact device model of either a BJT or HBT, the multiple-emitter thermal impedance model and the thermal coupling network form a complete electrothermal model suitable for DC, AC and transient device/circuit simulation. This type of electrothermal model is generally more efficient for circuit simulation than either finite difference or finite element techniques; however, it can be quite complex for devices with a large number of fingers and/or fingers with a large number of sub-cells. In such a case, simulating moderate- to large-size circuits could become impractical. The complexity of the electrothermal device model can be reduced by using a lumped modeling methodology. The measurement technique described by Zweidinger et al., referred to in this work as base-current thermometry, can extract the thermal impedance of a bipolar transistor using the temperature dependence of the base current [Zwe96]. By applying this extraction technique to the simulation of the complete electrothermal device model, a more compact lumped electrothermal can be produced. The lumped model implicitly contains all the details

of the thermal actions and interactions described by the complete electrothermal model, but with less complexity.

To present a clear discussion of the lumped electrothermal model generation methodology, a few definitions and conventions will be established. Due to the thermal interactions between fingers in a multiple-emitter device a lateral temperature gradient will exist across the device. Therefore, the current distribution among the fingers may not be uniform since the hotter fingers will carry a larger amount of current. As power dissipation increases, the lateral temperature gradient also increases and eventually the device will become unstable and enter either thermal runaway (BJT's) or current collapse (HBT's). Prior to the onset of thermal instability, the lateral thermal gradient is small and the current distribution among the fingers is approximately uniform. When the device reaches the point of thermal instability, the fingers no longer operate under similar bias conditions and the current no longer divides evenly among the fingers. Therefore, before a device becomes thermally unstable, it is defined to be in the uniform operating regime; and, once the device becomes unstable, it is defined to be in the nonuniform operating regime.

In the uniform operating regime, the EOT of the device varies linearly with the power and the complete electrothermal model can be represented by a single lumped device model and lumped thermal impedance, Z_{THL}. The circuit representation for the uniform lumped model is shown in Figure 3.4a; the emitter area of the lumped device model is equal to the total emitter area of the multiple-emitter device. As a device becomes thermally unstable the cooler fingers begin to turn off, leaving the hottest finger to conduct all of the current; the temperature-

(a)

(b)
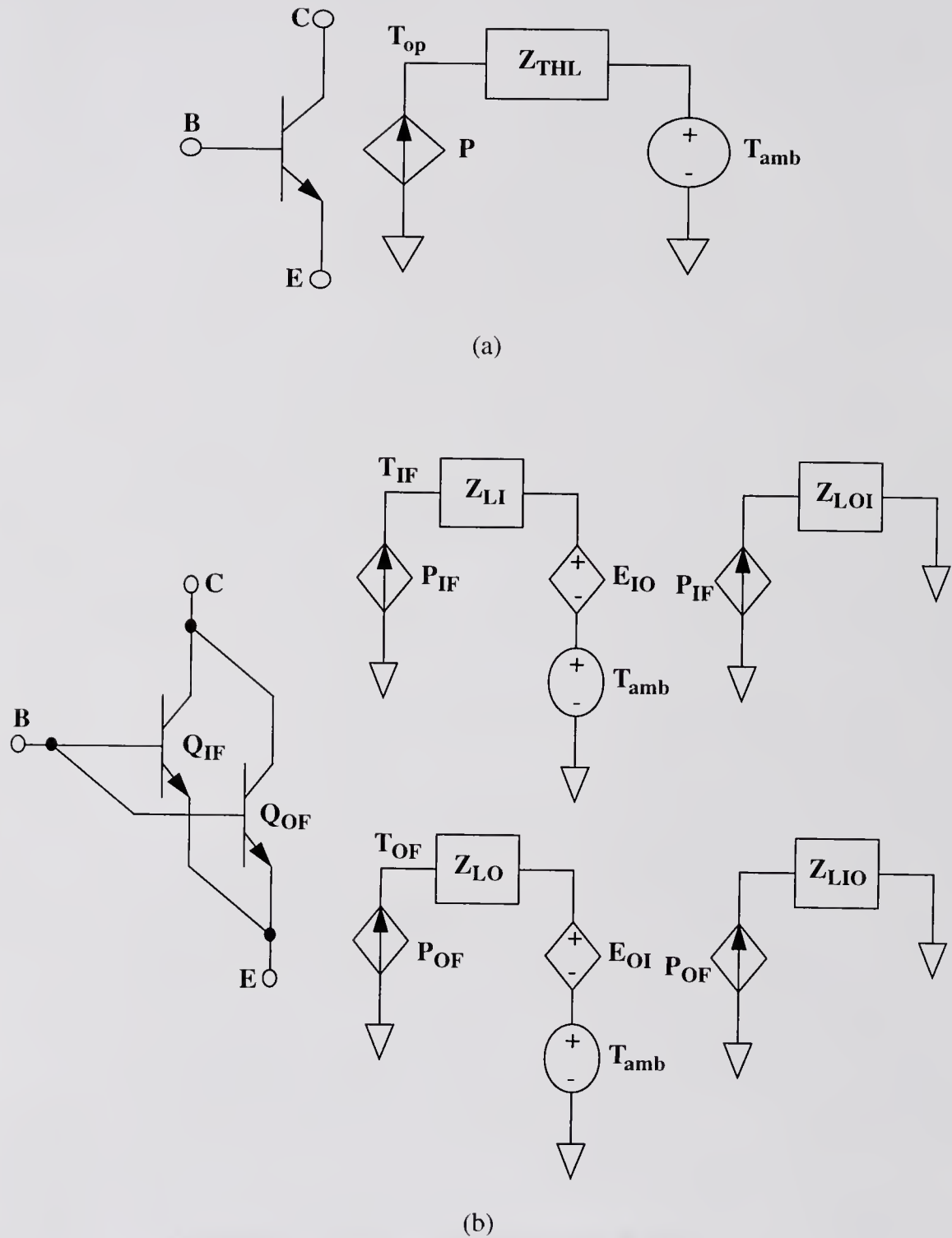
Figure 3.4   Circuit representations of the lumped multiple-emitter BJT/HBT electrothermal model: a) For the uniform operating regime; b) for the nonuniform operating regime.

power relation becomes nonlinear and the uniform lumped model will not accurately model the device characteristics. To model this shut-down mechanism, the nonuniform lumped model, which is shown in Figure 3.4b, uses two lumped device models and four lumped thermal impedances to represent the entire device. Device Q_{IF} is used to represent the hottest emitter finger. In a device with an odd number of fingers, the hottest finger will be the middle finger. If a device has an even number of emitter fingers, due to process variation, the hottest finger will be one of the inner most fingers. For consistency, in either case the hottest finger will be referred to as the middle finger. The other device model, Q_{OF}, represents the remaining outer emitter fingers. The emitter area of Q_{IF} is equal to that of a single emitter finger and the emitter area of Q_{OF} is equivalent to the sum of the emitter areas of the outer fingers. The lumped thermal impedances Z_{SI} and Z_{SO} model the self impedances of the middle finger and the outer fingers, respectively. In the case of Z_{SO}, the impedance represents the effective temperature rise in the lumped outer fingers due only to their power dissipation. The lumped coupling impedance Z_{CIO} models the temperature rise in the middle finger due to the power dissipation in the lumped outer fingers. The reciprocal coupling impedance Z_{COI}, corresponds to the effective temperature rise in the lumped outer fingers due to the power generation of the middle finger.

Thermally triggered instability in bipolar devices can lead to circuit failure and even catastrophic device failure. Typically, this region of operation is avoided in circuit design. Therefore, in most cases, the uniform lumped model should be appropriate for most applications. However, if the effects of thermal instability on

device/circuit operation need to be investigated the nonuniform model should be used.

3.3.1 A Review of Base-Current Thermometry

Base-current thermometry uses the base current as a thermometer to extract the thermal impedance of a bipolar transistor [Zwe96]. The technique was developed for measurement-based extraction but can be applied to the simulation of compact device models as long as the models' temperature dependences are physically valid.

The first step of the procedure is to determine the dependence of the base current on temperature. The response of the base current to changes in temperature is represented by the fractional temperature coefficient, defined as

$$TC_F(I_B) \equiv \frac{1}{I_B}\frac{\partial I_B}{\partial T}. \tag{3.5}$$

By biasing a device in the common-emitter configuration (avoiding impact-ionization), and separately varying the collector voltage and the ambient temperature, the thermal resistance of the device can be extracted Since the base-collector conductance is typically negligible, any changes in the base current during the measurements are due solely to the change in operating temperature. Therefore, once the self-heating effects are accounted for, the fractional temperature coefficient can be determined from the measured base current variations.

The second step of the procedure is to extract the transient thermal impedance. The collector and base currents of the device are monitored for a step in the collector voltage. The transient change in temperature can be expressed as

$$\Delta T(t) = \frac{I_B(t) - I_B(0)}{I_B TC_F(I_B)}, \tag{3.6}$$

where I_B is the median value of the base current for the transient. The temperature change is then normalized by the magnitude of the power step, giving the following equation for the thermal impedance

$$Z_{TH}(t) = \frac{\Delta T(t)}{\Delta P}. \tag{3.7}$$

3.3.2 Generation of the Lumped Electrothermal Model

The first step in the lumped model generation is to extract the temperature coefficient of the base current by performing DC SPICE simulations at different ambient temperatures. The .TEMP control card is used to set the ambient temperature; temperature steps between 4 and 10 degrees are sufficient, where a geometric mean can be used to average $TC_F(I_B)$ over temperature to correct for nonlinearities. The base voltage should be selected for the desired operating point and for each temperature setting, the collector voltage should be swept over a range in the forward active region. The range of collector voltages should be large enough to produce a linear increase in base current. Examples of the resulting base current characteristics are shown in Figure 3.5. The base and collector current values should
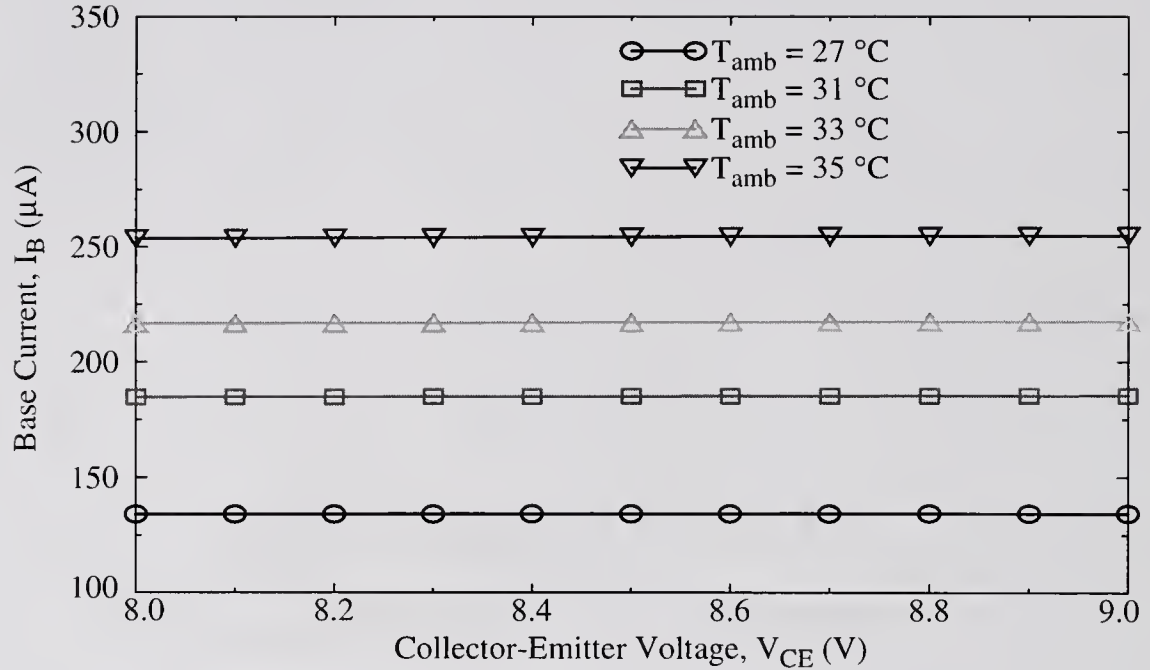
(a)
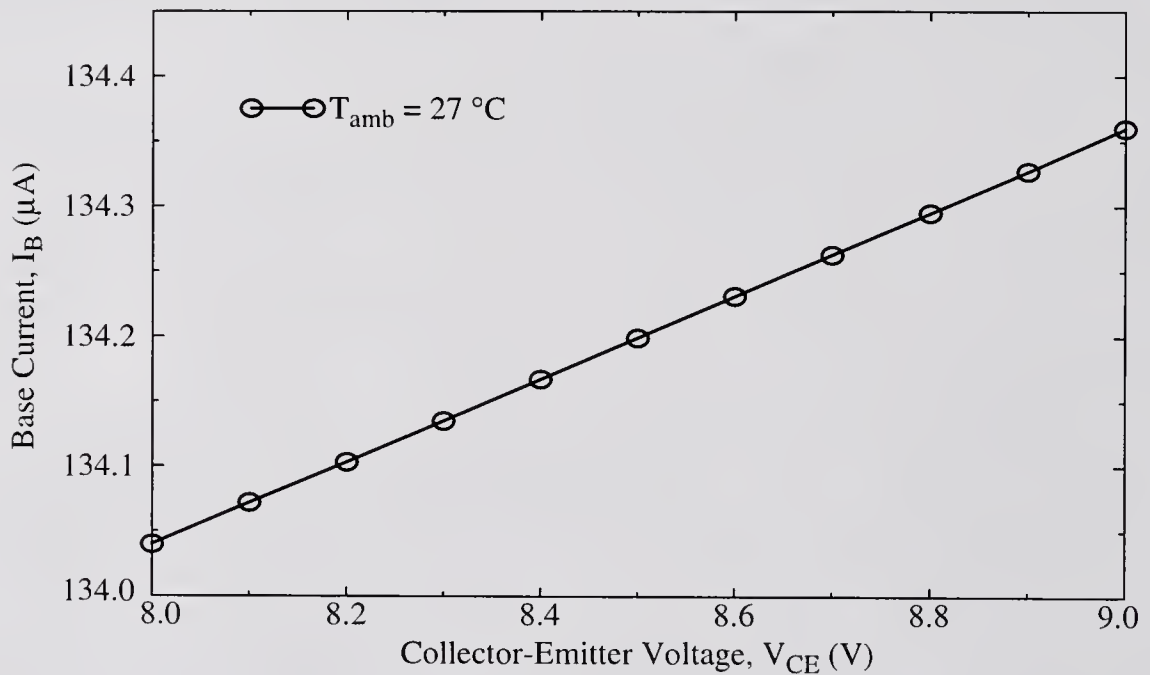
(b)

Figure 3.5   Simulated base current versus collector-emitter voltage: a) For varying
ambient temperatures; b) for T_{amb} = 27 °C.

be stored at each bias point and used to calculate $TC_F(I_B)$. For the uniform model, this process should be performed once, using the net base and collector currents of the complete device. For the nonuniform model, the procedure should also performed just once. The temperature rise in the middle finger, corresponding to either Z_{SI} or Z_{CIO}, can be taken directly from the multiple-emitter thermal impedance model, so an extraction is not necessary. Therefore, only the temperature coefficient of the lumped outer fingers should be needed. The coupling impedances from the middle finger to each of the outer fingers should be turned off and the net currents of only the lumped outer fingers are used to calculate $TC_F(I_B)$.

The second step of the procedure is to generate the transient thermal response. The transistor should be set in the common-emitter configuration with the base voltage set to the value used to extract $TC_F(I_B)$. The collector voltage should be stepped between two bias points in the forward-active region. The combination of the selected base voltage and collector voltage step-size should set the current level such that a significant base current response results from the step in power. The rise-time of the voltage step should be faster than the shortest expected thermal time constant and the step length should be long enough to allow the current response to reach steady-state. The base current should be recorded during the transient as well as the collector current values at the start and end of the step. Figure 3.6a shows an example of the transient base current response. Using (3.6) and the $TC_F(I_B)$ from the first step, the transient base current response can be converted into the transient temperature response. The magnitude of the power step can then used to normalize the temperature response, resulting in the transient thermal impedance; an example
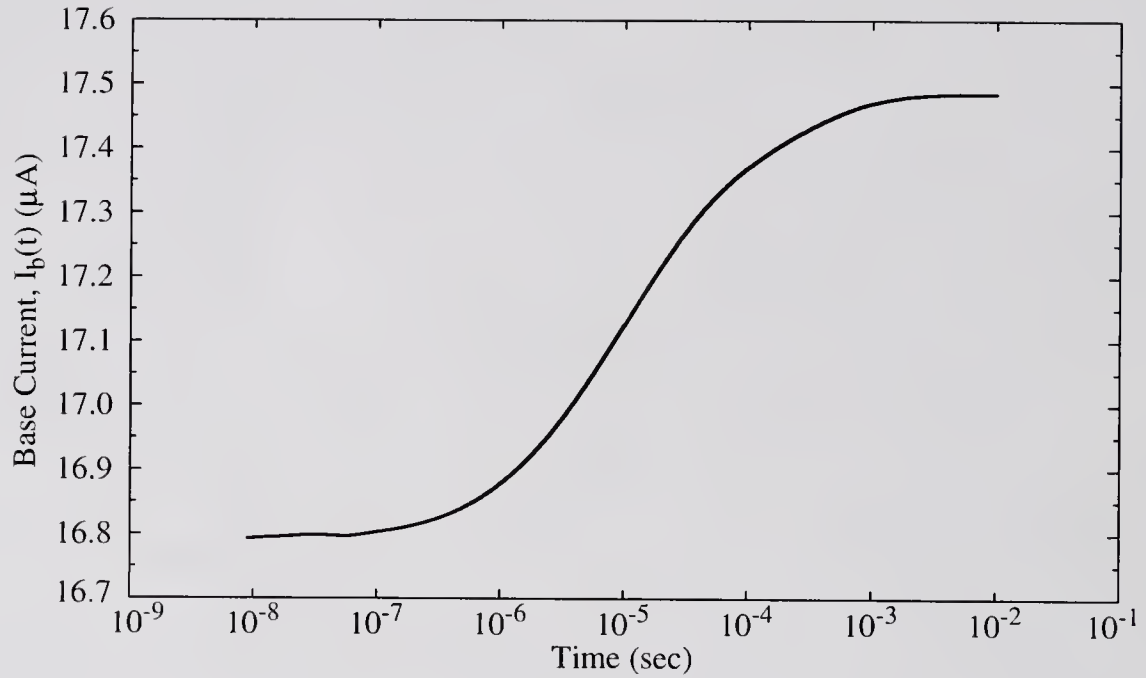
(a)
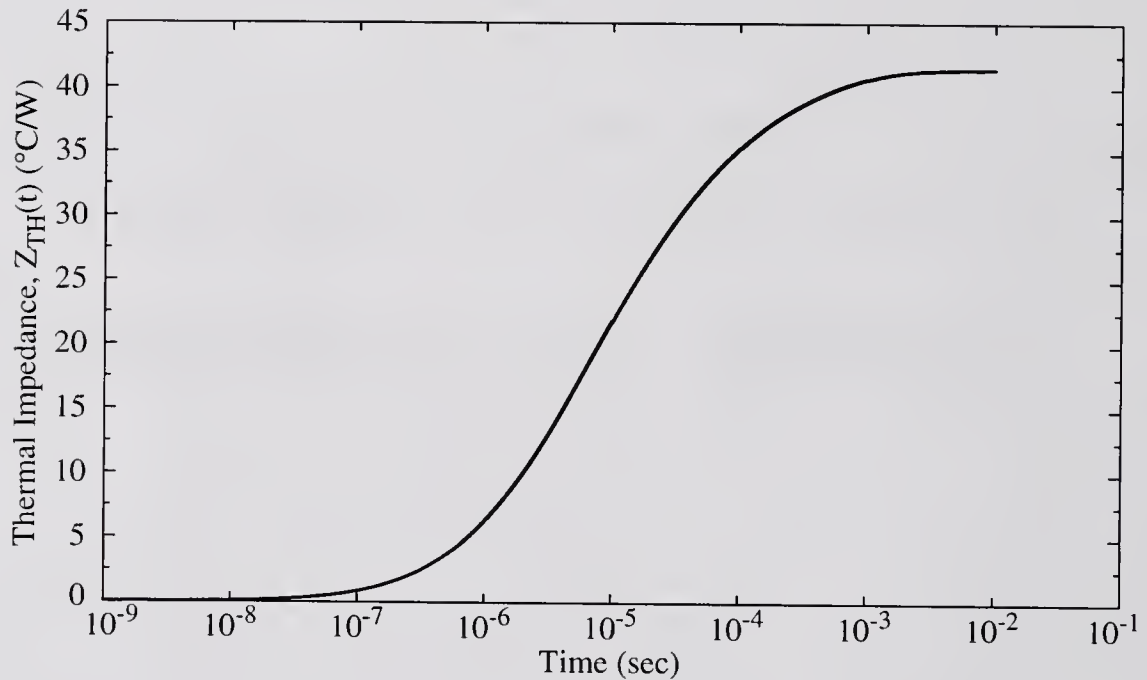
(b)

Figure 3.6   Simulated transient response of bipolar transistor to a step in collector voltage: a) Base current; b) temperature rise that is extracted from base current response.

is shown in Figure 3.6b. For the uniform lumped model, this procedure should be performed once, using the net transient base current response to calculate the lumped thermal impedance, Z_{THL}. For the nonuniform lumped model, the procedure should only be performed for the extraction of Z_{SO} and Z_{COI} from the net base current of the lumped outer fingers. When extracting Z_{SO}, the coupling impedances from the middle finger to the outer fingers should be turned off. The calculated temperature response should be normalized by the power dissipated in the lumped outer fingers. When extracting Z_{COI}, the self impedances and the coupling impedances between each of the outer fingers should be turned off. The calculated temperature response should be normalized by the power dissipated in the middle finger. The self impedance Z_{SI} in the lumped model corresponds to the self-impedance of a single finger, which can be calculated directly using the multiple-emitter thermal impedance model. The lumped coupling impedance Z_{CIO} can also be calculated directly from the multiple-emitter thermal impedance model; for a device with an odd number of fingers

$$Z_{CIO} = \frac{2}{(n-1)} \sum_{i=1}^{(n-1)/2} Z_{Ci} \tag{3.8}$$

and for a device with an even number of fingers

$$Z_{CIO} = \frac{1}{(n-1)} \left[Z_{C(n/2)} + 2 \sum_{i=1}^{(n-2)/2} Z_{Ci} \right] \tag{3.9}$$

where n is the total number of emitter fingers in the device.

3.4 Verification of the Lumped Electrothermal Model

To test the accuracy of the lumped electrothermal models, they are compared to the full electrothermal models using DC, AC and transient SPICE simulations. The following simulations of homojunction bipolar transistors use the QBBJT model in a modified version of MMSPICE [Jeo89, Lee96]. The HBT simulations were performed using a modified version of SPICE 2G.6 [Zwe97].

Figure 3.7 shows the simulated DC current characteristics for a ten-finger BJT. The device remains in the uniform operation regime for the simulated biases, and the lumped models produce nearly identical results. The discrepancies between the lumped and full electrothermal models are the greatest when the device is biased with a constant base current, in which case the error is less than 3%. When the device is biased with a fixed base voltage, the error between the lumped models and the full electrothermal models is no greater than 0.5%. Figure 3.8 shows the simulated DC current characteristics for a five-finger HBT. When the device is biased with a fixed base voltage, it remains in the uniform operation regime and the results are similar to those for the homojunction bipolar device. Both lumped models produce similar results and match the full electrothermal model to within 3% error. However, when the device is biased with a constant base current, it becomes thermally unstable and goes into current collapse. Under these operating conditions, the outer fingers shut down, leaving the middle finger to conduct all the current. The uniform lumped model does not show the collapse phenomenon which results in a 17% error at the highest bias point. The nonuniform lumped model, however, does a good job of representing the full electrothermal model, producing no more than 2% error.
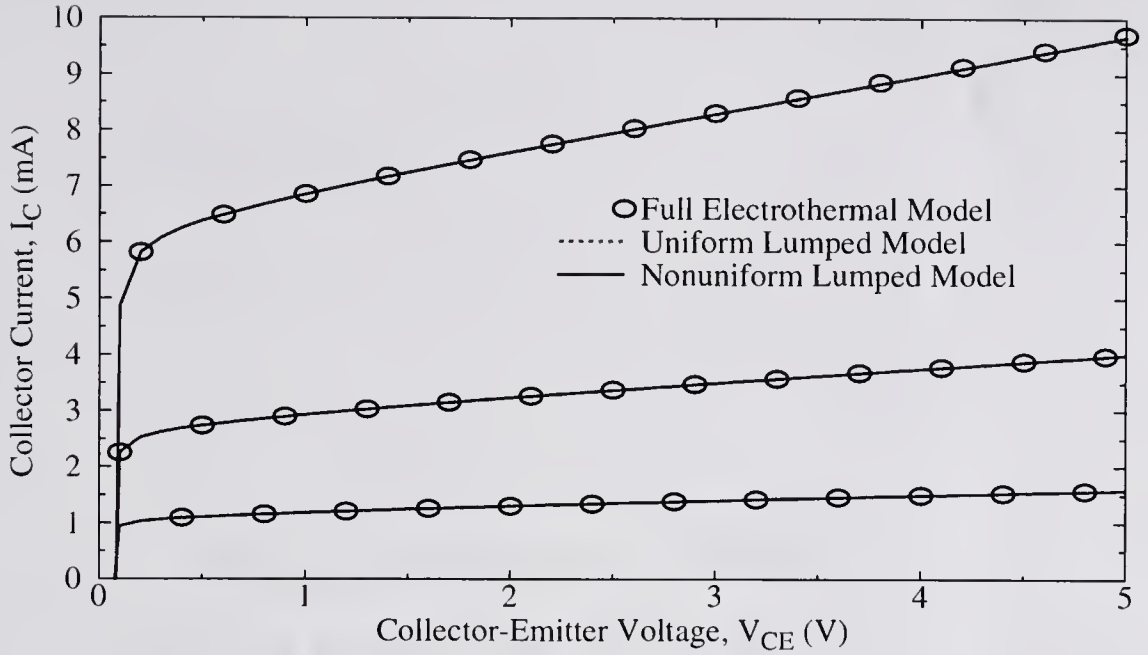
(a)
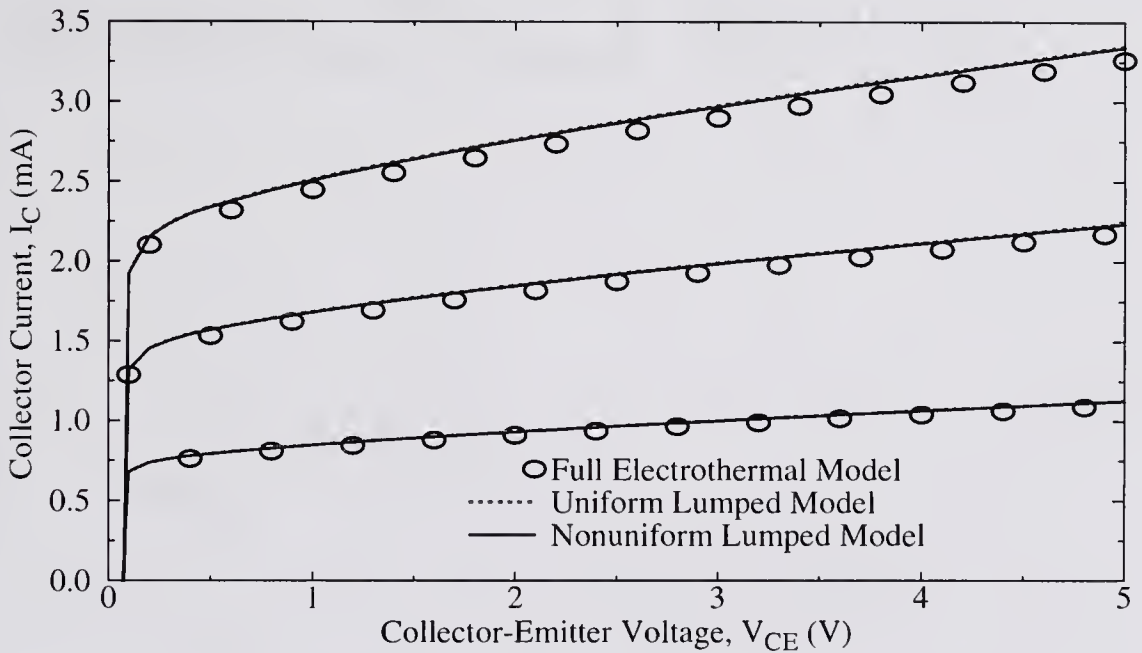
(b)

Figure 3.7   Simulated DC current characteristics for a ten-finger BJT with $A_E = 20 \times 1.6 \; \mu m^2$ for each emitter finger: a) The collector current as a function of collector emitter voltage for $V_{BE} = 0.75$, 0.775 and 0.80 V; b) the collector current as a function of collector emitter voltage for $I_B = 10$, 20 and $30 \; \mu A$.
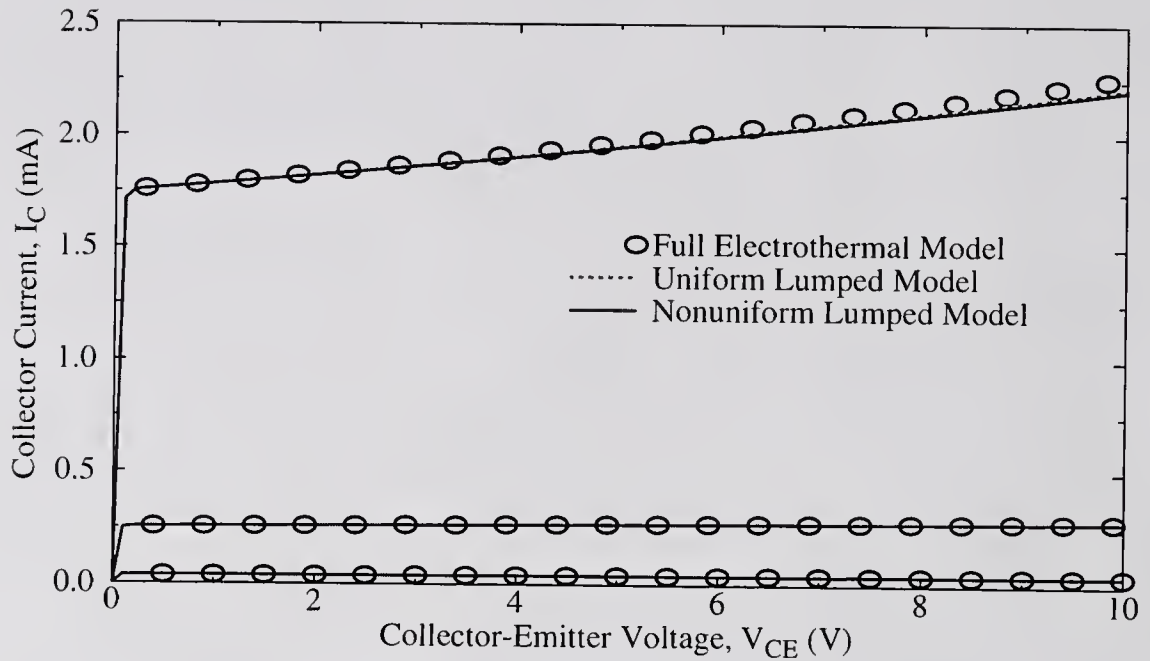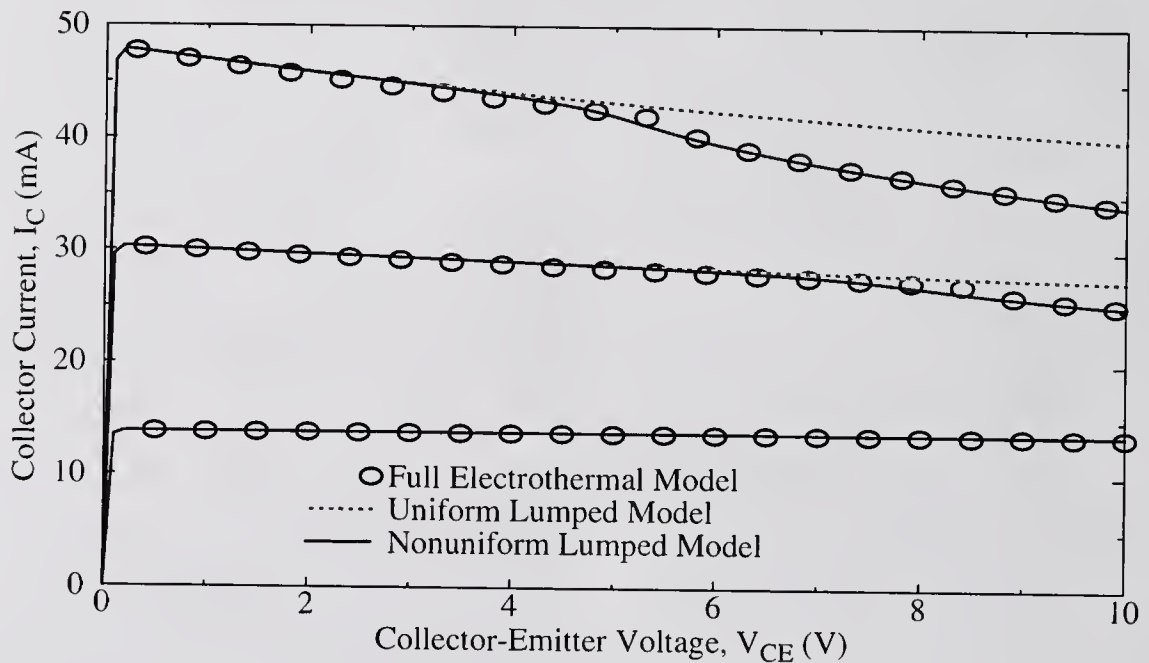
Figure 3.8   Simulated DC current characteristics for a five-finger HBT with $A_E = 20 \times 2 \ \mu m^2$ for each emitter finger: a) The collector current as a function of collector emitter voltage for $V_{BE} = 0.45$, 0.50 and 0.55 V; b) the collector current as a function of collector emitter voltage for $I_B = 1$, 2 and 3 mA.

The small-signal and transient performance of the lumped models were tested by simulating a cascode amplifier composed of two five-finger BJT's. The results of the simulations are shown in Figure 3.9. The lumped models produce almost identical results (the curves appear on top of each other) and agree closely with the full electrothermal model. The error in the small-signal gain is less than 1% and is due to errors in the DC operating point. Errors in the magnitude of the transient output waveform are less than 1%. For both the small-signal and transient simulations, the lumped models correctly reflect the frequency- and time-domain responses. The lumped models match the unity-gain frequency and phase response to within 4% of the full electrothermal model.

The benefit of the lumped models is that they have fewer components and less complexity than the full electrothermal model. The reduction in complexity results in models that are more efficient to simulate, at the cost of some accuracy. One portion of the increased simulation time of the full electrothermal model is due directly to the added components used for the thermal model. The other part is due to the increase in the number of iterations required to reach convergence in the presence of thermal feedback. Thus, the nonuniform lumped model does not offer as much of a speed enhancement over the full electrothermal model as the uniform lumped model. The total job times for simulations shown in this section (performed on a SPARCstation 5) are given in Table 3.1.
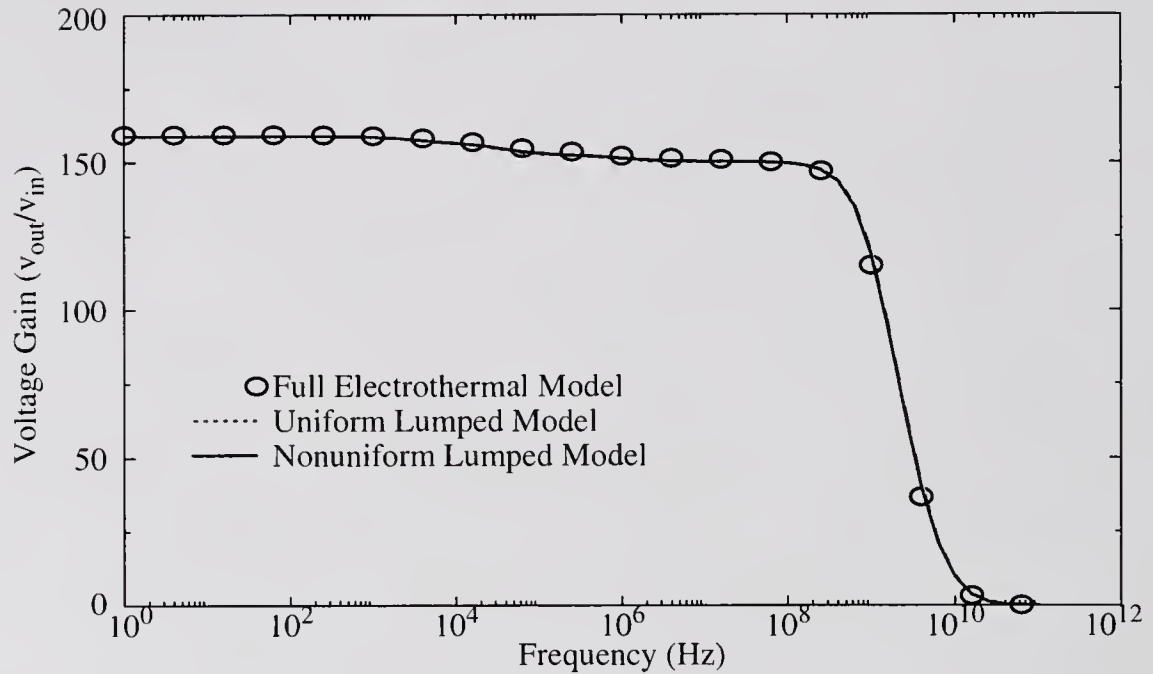
(a)
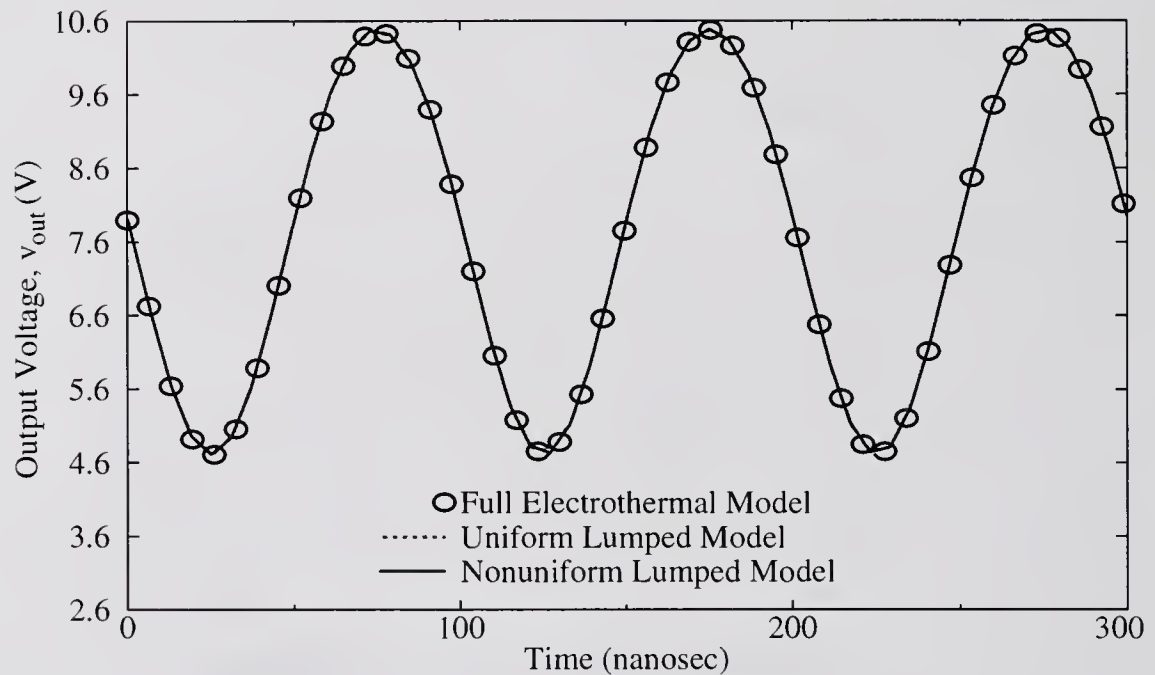
(b)

Figure 3.9   Small-signal and transient simulations of a cascode amplifier using five-finger BJT's with $A_E = 20 \times 1.6$ μm^2 for each emitter finger: a) Small-signal voltage gain; b) the output voltage resulting from a 40 mV peak-to-peak sinewave input at 10 MHz.

Table 3.1 Total simulation time

Simulation		Time (sec)		
		Full	Uniform	Nonuniform
HBT: 5 fingers	Fixed V_{BE}	20.05	1.13	3.82
	Fixed I_B	34.23	1.35	7.17
BJT: 10 fingers	Fixed V_{BE}	159.93	6.83	14.17
	Fixed I_B	180.30	8.27	18.00
Cascode Amp	Small-signal	28.25	0.97	3.17
	Transient	62.98	5.87	15.77

In general, the uniform lumped model provided a 91% to 97% speed increase over the full electrothermal model. The nonuniform lumped model, while not as efficient as the uniform model, was still 75% to 91% faster than the full electrothermal model.

3.5 Summary

A circuit model that correctly represents thermal coupling between devices was developed and demonstrated using DC, AC and transient circuit simulations. The thermal coupling circuit is used in conjunction with the multiple-emitter thermal impedance model, providing circuit simulators with the capability to perform electrothermal simulations of multiple-emitter bipolar devices. In addition, the thermal coupling model is also applicable to the general case of thermal coupling between devices in an arbitrary circuit. The complexity of the electrothermal

multiple-emitter device model was reduced by extracting a lumped thermal response. In most cases, a multiple-emitter device can be represented by a single compact device model and a lumped thermal impedance. The lumped model offers a significant speed increase over the full electrothermal model, resulting in more efficient circuit simulations.

CHAPTER 4
A THREE-DIMENSIONAL THERMAL IMPEDANCE MODEL FOR VERTICAL BIPOLAR TRANSISTORS FABRICATED WITH FULL DIELECTRIC ISOLATION

4.1 Introduction

Bipolar junction transistors fabricated using full dielectric isolation (DIBJT's) offer many advantages over those fabricated with junction isolation in bulk wafer technologies. The parasitic capacitances and leakage currents from the collector to the substrate and from the collector to the junction-isolation implant are reduced with full dielectric isolation, which enhances device speed [Dav92]. The area taken up by the lateral dielectric isolation is typically smaller than the area of the diffused junction-isolation implant; therefore, full dielectric isolation is a means to increase the transistor packing density [Jer93]. Finally, full dielectric isolation also negates latchup, and improves radiation hardness [Gan92].

Full dielectric isolation of vertical BJT's can be achieved by using silicon-on-insulator (SOI) substrates with trench isolation. Figure 4.1 shows a diagram of a typical npn DIBJT. Direct wafer bonding (DWB) has become a common way to fabricate SOI substrates for bipolar technologies since it produces high-quality, defect-free SOI films [Nis91, Dav92, Fei92, Jer93, Nak95]. The DWB technique thermally bonds one semiconductor wafer to the oxidized surface of another wafer. One wafer serves as the substrate and the other wafer is used for the actual device

Figure 4.1   Cross-section of a typical bipolar transistor fabricated with full dielectric isolation (DIBJT).

fabrication. The oxide layer interposed between the two wafers becomes the buried isolation layer and is typically 0.4 to 2.0 µm thick [Dav92, Fei92, Nak95]. The device-fabrication wafer is thinned using chemical/mechanical polishing (CMP). One disadvantage of DWB is that the resulting silicon film can have large variation in thickness [Nis91]. Once the SOI wafer is prepared, the BJT fabrication process can follow the typical steps used for bulk, trench-isolated devices. The back-fill for the trench isolation is usually formed with chemical vapor deposition (CVD) oxide (0.1 to 1.0 µm thick) or a combination of CVD oxide and polysilicon (0.5 to 2.0 µm thick) [Nis91, Fei92, Nak95]. The device region enclosed by the trench isolation is referred to as the "tub," and the region surrounding the trenches is referred to as the "exterior silicon." Typical tub thicknesses range from 1.5 to 10 µm [Nis91, Fei92, Nak95].

A major disadvantage of full dielectric isolation is an increase in self-heating. The oxide used in the trench and buried isolation has a low thermal conductivity and impedes the flow of heat away from the device, resulting in higher operating temperatures. The thermal resistance of a DIBJT can be three times larger than that of its bulk counterpart [Gan92]. Previous methods for determining the thermal resistance of DIBJT's have relied on measurement-based extraction or complex numerical techniques such as finite-element solutions [Nis91, Gan92]. Both of these approaches have been limited to steady-state operation and do not provide insight into the dynamic variation of temperature. Furthermore, neither approach is practical for use in circuit simulation: temperature measurements are complicated and are not predictive; finite-element solutions are predictive but can require large amounts of computing time.

This chapter details the derivation of a physics-based model for the dynamic thermal impedance of DIBJT's operating in the forward active region. The effects of interconnect metallization on the thermal impedance are then investigated. The model is compared to three-dimensional finite-element simulations and measurements for verification. The model derivation is then simplified for the limiting case of steady-state heat conduction, resulting in a single closed-form equation for the thermal resistance of DIBJT's. The limitations of the thermal resistance model are shown with comparisons to measurements and to the full thermal impedance model.

4.2 Derivation of the DIBJT Thermal Impedance Model

For the derivation of the DIBJT thermal impedance model, the silicon tub is represented by a homogeneous finite medium with an adiabatic top surface (no heat transfer perpendicular to the surface). The interface between the buried oxide and the substrate is assumed to be at a uniform temperature, T_0. Figure 4.2 shows a three-dimensional cross-section of the simplified device geometry assumed for the model derivation. The definitions of the model parameters are given in Table 4.1. The imbedded heat source represents the base/collector space-charge region (SCR), and is modeled by a rectangular volume. The thickness of the base/collector SCR can be estimated using the depletion approximation as given in Chapter Two. The heat generated in this region is assumed to be due to uniform power dissipation. As verified for the bulk BJT thermal impedance model, this assumption is reasonable for
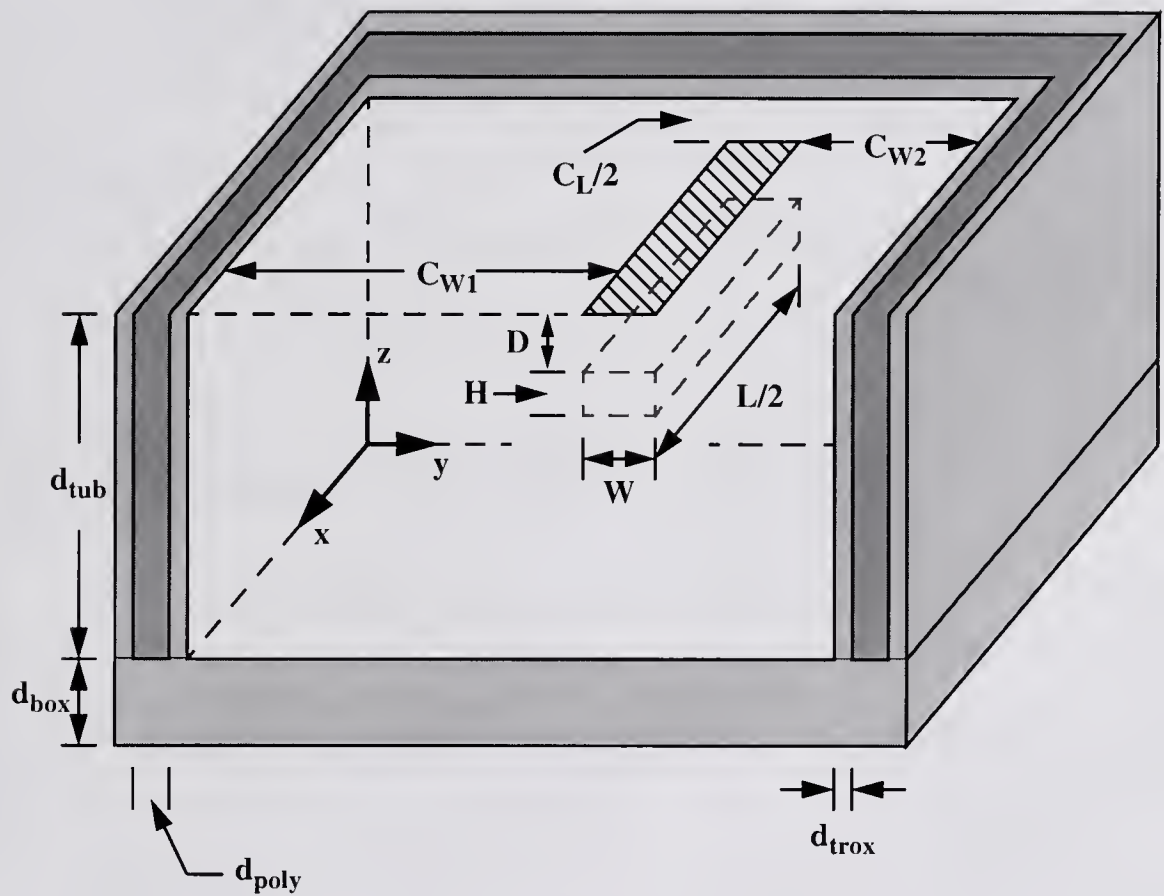
Figure 4.2   Cross-section of the simplified device geometry used to define the solution domain for the DIBJT thermal impedance model. The silicon tub is represented by a homogeneous finite medium with an adiabatic top surface. The model parameters are defined in Table 4.1.

Table 4.1 DIBJT thermal impedance model parameters

Parameter	Definition
W	Width of emitter stripe
L	Length of emitter stripe
D	Depth of base/collector junction
H	Thickness of base/collector SCR
d_{poly}	Thickness of trench polysilicon
d_{trox}	Thickness of trench oxide
d_{tub}	Thickness of silicon tub
d_{box}	Thickness of buried oxide
W_{tub}	Width of silicon tub
L_{tub}	Length of silicon tub
C_{W1}	Scaling constant for tub width
C_{W2}	Scaling constant for tub width
C_L	Scaling constant for tub length
θ_W	Spreading angle of heat flow
θ_L	Spreading angle of heat flow

the forward-active region before the onset of high-current effects. The heat source is displaced beneath the surface of the wafer by a distance assumed to be the depth of the base/collector junction. Any encroachment of the base/collector SCR into the base region is neglected since the base typically has a higher doping than the epi collector. The width and length of the silicon tub are assumed to scale directly with the width and length of the emitter stripe by the relations

$$W_{tub} = W + C_{W1} + C_{W2} \qquad (4.1)$$

and

$$L_{tub} = L + C_L, \qquad (4.2)$$

where C_{W1}, C_{W2} and C_L are constants that depend on the fabrication process.

Since the tub material is assumed to be homogeneous, the model neglects the effects of the LOCOS isolation that is used to cap the trench structure and to separate the base and emitter from the collector contact implant. This assumption is reasonable since the LOCOS is typically shallow semi-recessed LOCOS. The adiabatic boundary condition at the top surface of the device dictates that conduction through the interconnects and conduction/convection from the surface are neglected. This assumption is valid for the device regions that lie under thick silicon dioxide, since the thermal conductance from the device to the overlying oxide is approximately two to three orders of magnitude smaller than the device-to-substrate conductance. The effects of conduction via the interconnect metallization are examined later in the chapter.

The temperature rise at any point within the tub can be described by the nonhomogeneous three-dimensional heat conduction equation

$$\nabla^2 \Delta T_{tub}(x, y, z, t) + \frac{g(x, y, z, t)}{k_{si}} = \frac{1}{\alpha_{si}} \frac{\partial \Delta T_{tub}(x, y, z, t)}{\partial t} \tag{4.3}$$

and the following boundary conditions

$$\left[\frac{\partial \Delta T_{tub}(x, y, z, t)}{\partial x} - H_{tr}(t) \cdot \Delta T_{tub}(x, y, z, t) \right]\Bigg|_{x = 0} = 0 \tag{4.4}$$

$$\left[-\frac{\partial \Delta T_{tub}(x, y, z, t)}{\partial x} - H_{tr}(t) \cdot \Delta T_{tub}(x, y, z, t) \right]\Bigg|_{x = L_{tub}} = 0 \tag{4.5}$$

$$\left[\frac{\partial \Delta T_{tub}(x, y, z, t)}{\partial y} - H_{tr}(t) \cdot \Delta T_{tub}(x, y, z, t) \right]\Bigg|_{y = 0} = 0 \tag{4.6}$$

$$\left[-\frac{\partial \Delta T_{tub}(x, y, z, t)}{\partial y} - H_{tr}(t) \cdot \Delta T_{tub}(x, y, z, t) \right]\Bigg|_{y = W_{tub}} = 0 \tag{4.7}$$

$$\left[-\frac{\partial \Delta T_{tub}(x, y, z, t)}{\partial z} + H_{box}(t) \cdot \Delta T_{tub}(x, y, z, t) \right]\Bigg|_{z = 0} = 0 \tag{4.8}$$

$$\frac{\partial \Delta T_{tub}(x, y, z, t)}{\partial z}\Bigg|_{z = d_{tub}} = 0 \tag{4.9}$$

where ΔT_{tub} is the temperature rise above the local reference temperature $(\Delta T_{tub} = T_{tub} - T_0)$, g is the internal energy generation density, k is thermal

conductivity, α is thermal diffusivity ($\alpha = k/(\rho \cdot c_p)$ where ρ is density and c_p is specific heat) and t is time. Typical values for the material properties are given in Table 4.2. The terms $H_{tr}(t)$ and $H_{box}(t)$ are normalized heat-transfer coefficients that model the time-dependent heat flow through the trench and buried oxide, respectively. Equations (4.3) through (4.9) assume that the thermal conductivity is independent of temperature and position. For the tub material, the variation of k with temperature can be accounted for with the Kirchoff transformation detailed in Chapter Eight. However, the thermal impedance of a DIBJT is mainly determined by the isolation structures, and the thermal conductivity of silicon dioxide varies by less than 13% from 303 to 433 K [Goo95]. Also, as shown in Chapter Two, the thermal conductivity remains approximately constant over a wide impurity doping range, and any spatial dependence of the thermal conductivity can be neglected.

With the initial temperature rise of the device specified as

$$\Delta T_{tub}(x, y, z, 0) = 0, \tag{4.10}$$

the solution to (4.3) can be expressed in the form

$$\Delta T_{tub}(x, y, z, t) = \frac{\alpha_{si}}{k_{si}} \int_{t'=0}^{t} dt' \int_{V} G(x, y, z, t | x', y', z', t')g(x', y', z', t')dv' \tag{4.11}$$

where

Table 4.2 Material Properties

Property	Definition	Value
k_{si}	Thermal conductivity of silicon	1.412 (W cm^{-1} K^{-1})
ρ_{si}	Density of silicon	2.328 (g cm^{-3})
c_{psi}	Specific heat of silicon	0.7 (J g^{-1} K^{-1})
k_{ox}	Thermal conductivity of silicon dioxide	0.014 (W cm^{-1} K^{-1})
ρ_{ox}	Density of silicon dioxide	2.19 (g cm^{-3})
c_{pox}	Specific heat of silicon dioxide	1.4 (J g^{-1} K^{-1})
k_{poly}	Thermal conductivity of polysilicon	1.412[‡] (W cm^{-1} K^{-1})
ρ_{poly}	Density of polysilicon	2.328[‡] (g cm^{-3})
c_{ppoly}	Specific heat of polysilicon	0.7[‡] (J g^{-1} K^{-1})

Source: [Mul77]

[‡] Assumed to be equivalent to bulk silicon

$$G(x, y, z, t | x', y', z', t') = \left[\sum_{m=1}^{\infty} \exp[-\alpha_{si}\beta_m^2(t-t')] \frac{1}{N(\beta_m)} X(\beta_m, x) X(\beta_m, x') \right]$$

$$\cdot \left[\sum_{n=1}^{\infty} \exp[-\alpha_{si}\gamma_n^2(t-t')] \frac{1}{N(\gamma_n)} Y(\gamma_n, y) Y(\gamma_n, y') \right]$$

$$\cdot \left[\sum_{p=1}^{\infty} \exp[-\alpha_{si}\eta_p^2(t-t')] \frac{1}{N(\eta_p)} Z(\eta_p, z) Z(\eta_p, z') \right] \quad (4.12)$$

is the Green's function for the given boundary-value problem [Ozi93]. The expressions for the eigenfunctions are determined by the boundary conditions for each direction. For the x- and y-directions

$$X(\beta_m, x) = \beta_m \cos(\beta_m x) + H_{tr}(t) \cdot \sin(\beta_m x) \quad (4.13)$$

$$Y(\gamma_n, x) = \gamma_n \cos(\gamma_n x) + H_{tr}(t) \cdot \sin(\gamma_n x) \quad (4.14)$$

$$\frac{1}{N(\beta_m)} = 2[W_{tub}\{\beta_m^2 + H_{tr}^2(t)\} + 2H_{tr}(t)]^{-1} \quad (4.15)$$

$$\frac{1}{N(\gamma_n)} = 2[L_{tub}\{\gamma_n^2 + H_{tr}^2(t)\} + 2H_{tr}(t)]^{-1} \quad (4.16)$$

where the eigenvalues are determined by the positive roots of the following transcendental equations

$$\tan(\beta_m L_{tub}) = \frac{2\beta_m H_{tr}(t)}{\beta_m^2 - H_{tr}^2(t)} \quad (4.17)$$

$$\tan(\gamma_n W_{tub}) = \frac{2\gamma_n H_{tr}(t)}{\gamma_n^2 - H_{tr}^2(t)} \ . \tag{4.18}$$

For the z-direction

$$Z(\eta_p, z) = \cos[\eta_p(d_{tub} - z)] \tag{4.19}$$

$$\frac{1}{N(\eta_p)} = \frac{2[\eta_p^2 + H_{box}^2(t)]}{d_{tub}[\eta_p^2 + H_{box}^2(t)] + H_{box}(t)} \tag{4.20}$$

where the eigenvalues are determined by the positive roots of the following transcendental equation

$$\eta_p \tan(\eta_p d_{tub}) = H_{box}(t) \ . \tag{4.21}$$

Equation (4.12) physically represents the temperature rise at any point (x, y, z) in the tub at time t, due to an instantaneous point source at point (x', y', z') at time t'. To account for the heat-generation volume $(V = W \cdot L \cdot H)$, (4.12) is integrated over the base/collector SCR. Assuming a unit step increase in power dissipation at time $t' = 0$ and expressing the temperature rise in the tub as

$$\Delta T_{tub}(t) = Z_{TH}(t) \cdot P \tag{4.22}$$

yields the transient thermal impedance

$$Z_{TH}(x, y, z, t) = \frac{1}{\rho_{si}c_{psi}V} \int_t \left[\sum_{m=1}^{\infty} \exp[-\alpha_{si}\beta_m^2 t] \frac{1}{N(\beta_m)} X(\beta_m, x) X'(\beta_m) \right]$$

$$\cdot \left[\sum_{n=1}^{\infty} \exp[-\alpha_{si}\gamma_n^2 t] \frac{1}{N(\gamma_n)} Y(\gamma_n, y) Y'(\gamma_n) \right]$$

$$\cdot \left[\sum_{p=1}^{\infty} \exp[-\alpha_{si}\eta_p^2 t] \frac{1}{N(\eta_p)} Z(\eta_p, z) Z'(\eta_p) \right] dt \qquad (4.23)$$

where

$$X'(\beta_m) = 2\sin\left[\frac{\beta_m L}{2}\right] \left\{ \cos\left[\beta_m\left(\frac{C_L + L}{2}\right)\right] + \frac{H_{tr}(t)}{\beta_m} \sin\left[\beta_m\left(\frac{C_L + L}{2}\right)\right] \right\} \qquad (4.24)$$

$$Y'(\gamma_n) = 2\sin\left[\frac{\gamma_n W}{2}\right] \left\{ \cos\left[\gamma_n\left(C_{W1} + \frac{W}{2}\right)\right] + \frac{H_{tr}(t)}{\gamma_n} \sin\left[\gamma_n\left(C_{W1} + \frac{W}{2}\right)\right] \right\} \qquad (4.25)$$

$$Z'(\eta_p) = \frac{1}{\eta_p} \{ \sin[\eta_p(D + H)] - \sin[\eta_p D] \} . \qquad (4.26)$$

To represent the temperature rise in the device by a single effective value, (4.23) should be evaluated at a single point. The coordinates for surface corner of the emitter, $(x = C_L/2, y = C_{W1}, z = d_{tub})$, are substituted into (4.13), (4.14) and (4.19) to keep the model consistent with the bulk BJT thermal impedance model. As shown later in the chapter, the model evaluated at those coordinates agrees well with measured values of the steady-state thermal resistance.

Most of the parameters for the DIBJT thermal impedance model are fixed by the geometry of the device structure. However, the parameter H depends on the

electrical bias of the device and can change with operating conditions. As shown in Chapter Two, the base/collector SCR thickness, and hence the thermal impedance model, depends only moderately on the bias of the base/collector junction. For the DIBJT thermal impedance model, this dependence is weaker than that for the bulk BJT model since the dielectric isolation primarily determines the thermal impedance. The variation in the predicted thermal resistance of a DIBJT due to changes in H is approximately one-half to one-third that of a bulk BJT. The variation will increase as the thickness of the dielectric isolation is decreased or the tub scaling constants are increased; however, for most practical DIBJT structures, the dependence of the thermal impedance on bias can be neglected.

The DIBJT thermal impedance model can be extended to account for BJT's with multiple emitter fingers by integrating (4.12) over each base/collector SCR. Using an analysis similar to that in Chapter Two for bulk MEBJT's, expressions for the self and coupling impedances can be derived.

4.2.1 Derivation of the Buried-Oxide Heat-Transfer Coefficient

The normalized heat-transfer coefficient of the buried oxide, $H_{box}(t)$, describes the time-dependent heat conduction through the buried oxide to the substrate. The heat flux through the buried oxide is assumed to be predominantly one-dimensional (1-D). This assumption is valid for most of the tub area ($A_{tub} = W_{tub}L_{tub}$) but is questionable at the edges of the tub, where the heat flows out laterally under the trench. However, analogous to field fringing effects in parallel-plate capacitors, the proportion of lateral heat flow becomes smaller as the

tub area becomes larger. Figure 4.3 shows the results of ANSYS simulations of the heat flow through the buried oxide. The data corresponds to the z-component of the heat flux vector in the buried oxide, normalized by the heat flux vector magnitude (referred to as the "flux ratio"). The plots clearly show that the heat flux in the buried oxide is predominantly 1-D, even at the interface between the trench and the buried oxide.

The temperature rise in the buried oxide, $\Delta T_{ox} = T_{ox} - T_0$, can be described by the heat conduction equation

$$\frac{\partial^2 \Delta T_{box}}{\partial z^2} = \frac{1}{\alpha_{ox}} \frac{\partial \Delta T_{box}}{\partial t}. \tag{4.27}$$

The following boundary conditions are imposed at the interface between the tub and the buried oxide

$$k_{ox} \frac{\partial \Delta T_{ox}}{\partial z} = k_{si} \frac{\partial \Delta T_{tub}}{\partial z} \tag{4.28}$$

$$\Delta T_{ox} = \Delta T_{tub} \tag{4.29}$$

and at the interface between the buried oxide and the substrate

$$\Delta T_{ox} = 0. \tag{4.30}$$

Using the variable substitution
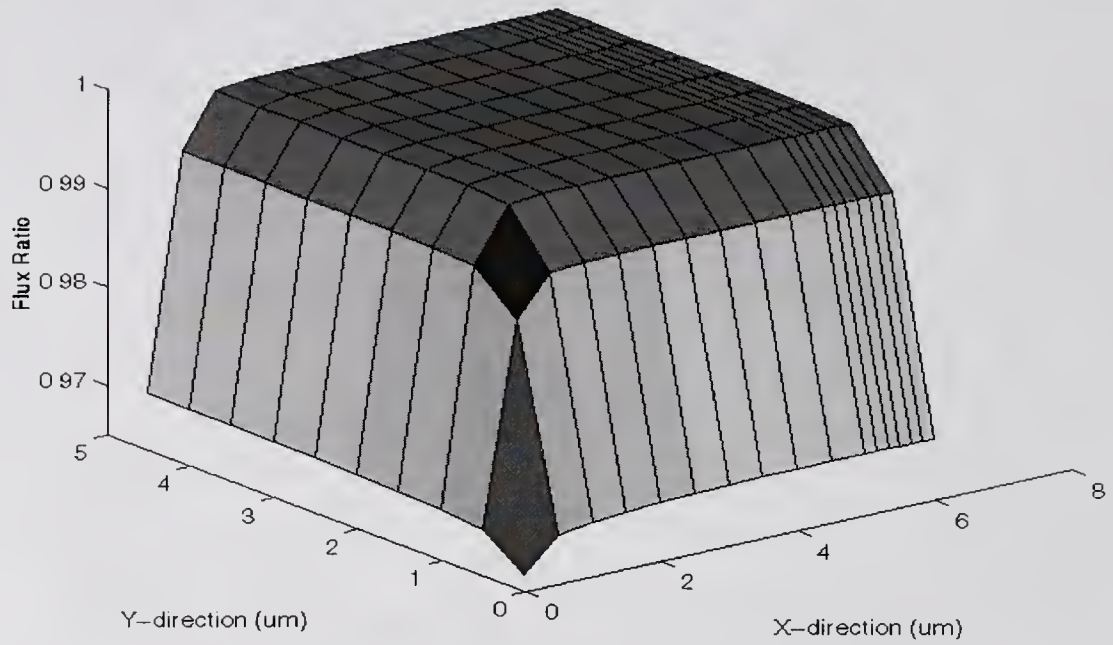
(a)
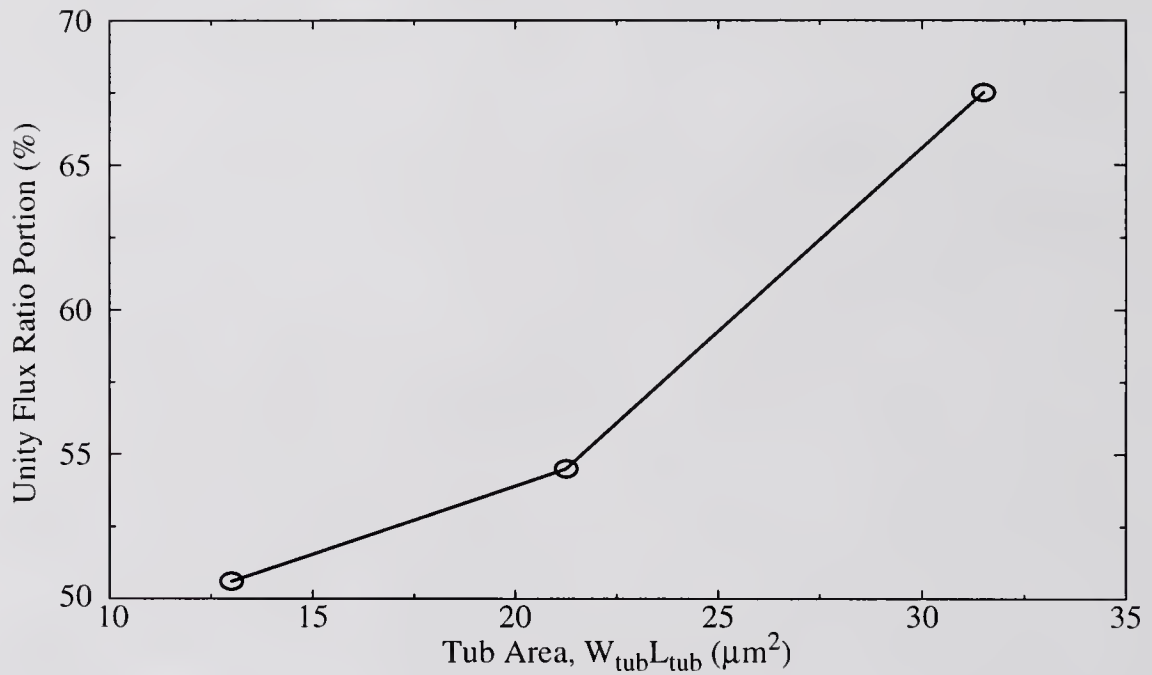
(b)

Figure 4.3   The ratio of the z-component of the heat flux vector to the magnitude of the total heat flux vector in the buried oxide with d_{box} = 1 μm. L = 2 μm and W = 0.5 μm: a) For C_{W1} = C_{W2} = $C_L/2$ = 5 μm; b) the portion of the tub area where the flux ratio is unity (signifying complete 1-D heat flow) for C_{W1} = C_{W2} = $C_L/2$ = 3, 4, and 5 μm.

$$\xi = \frac{x}{\sqrt{\alpha_{ox}t}},$$
(4.31)

(4.27) is transformed into the following ordinary differential equation

$$\frac{d^2\Delta T_{ox}}{d\xi^2} + \frac{\xi}{2}\frac{d\Delta T_{ox}}{d\xi} = 0.$$
(4.32)

Equation (4.32) has a general solution of the form

$$\Delta T_{ox} = c_1 \text{erf}\left(\frac{\xi}{2}\right) + c_2$$
(4.33)

where c_1 and c_2 are arbitrary constants. The equation that describes the flux at the interface between the tub and the buried oxide can be derived by substituting (4.31) into (4.33) and then applying the boundary conditions given by (4.28) through (4.30). The resulting expression

$$-k_{si}\frac{\partial \Delta T_{tub}}{\partial z} = \frac{k_{ox}\Delta T_{tub}}{\sqrt{\pi\alpha_{ox}t}\left[\text{erf}\left(\frac{d_{box}}{\sqrt{4\alpha_{ox}t}}\right)\right]}$$
(4.34)

when rearranged into the form given by (4.8), yields the normalized heat-transfer coefficient for the buried oxide

$$H_{ox}(t) = \frac{k_{ox}}{k_{si}\sqrt{\pi\alpha_{ox}t}\left[\text{erf}\left(\frac{d_{box}}{\sqrt{4\alpha_{ox}t}}\right)\right]}.$$
(4.35)

4.2.2 Derivation of the Trench Heat-Transfer Coefficient

The normalized heat-transfer coefficient of the trench isolation, $H_{tr}(t)$, describes the time-dependent heat conduction through the trench structure and exterior silicon to the substrate. To simplify the analysis, the thickness of the trench is assumed to be uniform along the z-direction and the composite trench structure is represented by a single material with the lumped thermal properties [Man90]

$$k_{tr} = \frac{d_{tr}k_{ox}k_{poly}}{2d_{trox}k_{poly} + d_{poly}k_{ox}} \tag{4.36}$$

$$d_{tr}\rho_{tr}c_{ptr} = 2d_{trox}\rho_{ox}c_{pox} + d_{poly}\rho_{poly}c_{ppoly} \tag{4.37}$$

$$d_{tr} = 2d_{trox} + d_{poly} \tag{4.38}$$

$$\alpha_{tr} = \frac{k_{tr}}{\rho_{tr}c_{ptr}}. \tag{4.39}$$

The heat flux through the trench is assumed to be predominantly one-dimensional (1-D). This assumption is valid for most of the trench area ($A_{tr} = W_{tub}d_{tub}$ or $L_{tub}d_{tub}$), but is questionable at the corners of the trench structure and at the edges between the trench walls and the buried oxide. Figure 4.4 shows the results of ANSYS simulations of the heat flow through the trench. The data corresponds to either the x- or the y-component of the heat flux vector in the trench which is normalized by the heat flux vector magnitude (referred to as the "flux ratio"). The plots clearly show

Figure 4.4   The ratio of the y-component of the heat flux vector to the magnitude of the total heat flux vector in the trench for half of the trench wall. The device parameters are $L = 2 \, \mu m$, $C_L = 10 \, \mu m$, $d_{trox} = 0.25 \, \mu m$ and $d_{poly} = 1.6 \, \mu m$.

that the heat flux in the trench is predominantly 1-D, even at the corners of the trench and at the interface between the trench and the buried oxide.

The temperature rise in the trench, $\Delta T_{tr} = T_{tr} - T_0$, can be described by the heat conduction equation

$$\frac{\partial^2 \Delta T_{tr}}{\partial n^2} = \frac{1}{\alpha_{tr}} \frac{\partial \Delta T_{tr}}{\partial t},$$

(4.40)

where n is either the x- or y-direction. The following boundary conditions are imposed at the interface between the tub and the trench

$$k_{tr} \frac{\partial \Delta T_{tr}}{\partial x} = k_{si} \frac{\partial \Delta T_{tub}}{\partial x}$$

(4.41)

$$\Delta T_{tr} = \Delta T_{tub}$$

(4.42)

and at the interface between the trench and the exterior silicon

$$k_{tr} \frac{\partial \Delta T_{tr}}{\partial n} = k_{si} \frac{\partial \Delta T_{si}}{\partial n}$$

(4.43)

$$\Delta T_{tr} = \Delta T_{si}.$$

(4.44)

The boundary conditions given by (4.43) and (4.44) require a solution for the temperature rise in the exterior silicon, $\Delta T_{si} = T_{si} - T_0$. The temperature in the exterior silicon can be derived by assuming that the heat flow is primarily in the directions normal to the trench walls. Therefore, the temperature rise in the exterior

silicon is assumed to be one-dimensional. The validity of such an assumption can be evaluated using the Biot number, which corresponds to the ratio of the internal and external thermal resistances of a given object [Ozi93]. If the Biot number for an object is much less than unity, then it can be approximated by a one-dimensional thermal medium. The vertical Biot number for the exterior silicon is $B_{Vsi} = (k_{ox}d_{tub})/(k_{si}d_{box})$, so that for $d_{tub} \ll 100 \cdot d_{box}$ the vertical temperature gradient can be neglected. The extent of the temperature gradients in the directions parallel to the trench walls is also investigated by simulating the heat flux in the exterior silicon at the trench/exterior silicon interface. Figure 4.5 shows the results of one such ANSYS simulation. For most of the trench area, the heat flux is primarily perpendicular to the trench walls. However, the one-dimensional assumption does break down around the corners of the trench where the heat flow becomes more two-dimensional. Therefore, the one-dimensional model will tend to slightly under predict the heat transfer through the trench.

The exterior silicon can be divided into four regions surrounding the trench; each region is modeled as a one-dimensional cooling fin. To approximately account for the lateral spread of heat around the corners of the trench, each cooling fin is assumed to have an increasing cross-sectional area. The temperature rise in the each exterior silicon cooling fin can be described by

$$\frac{\partial^2 \Delta T_{si}}{\partial n^2} + \frac{1}{(n + C_n)} \frac{\partial \Delta T_{si}}{\partial n} - m_{si}^2 \Delta T_{si} = \frac{1}{\alpha_{si}} \frac{\partial \Delta T_{si}}{\partial t}. \tag{4.45}$$
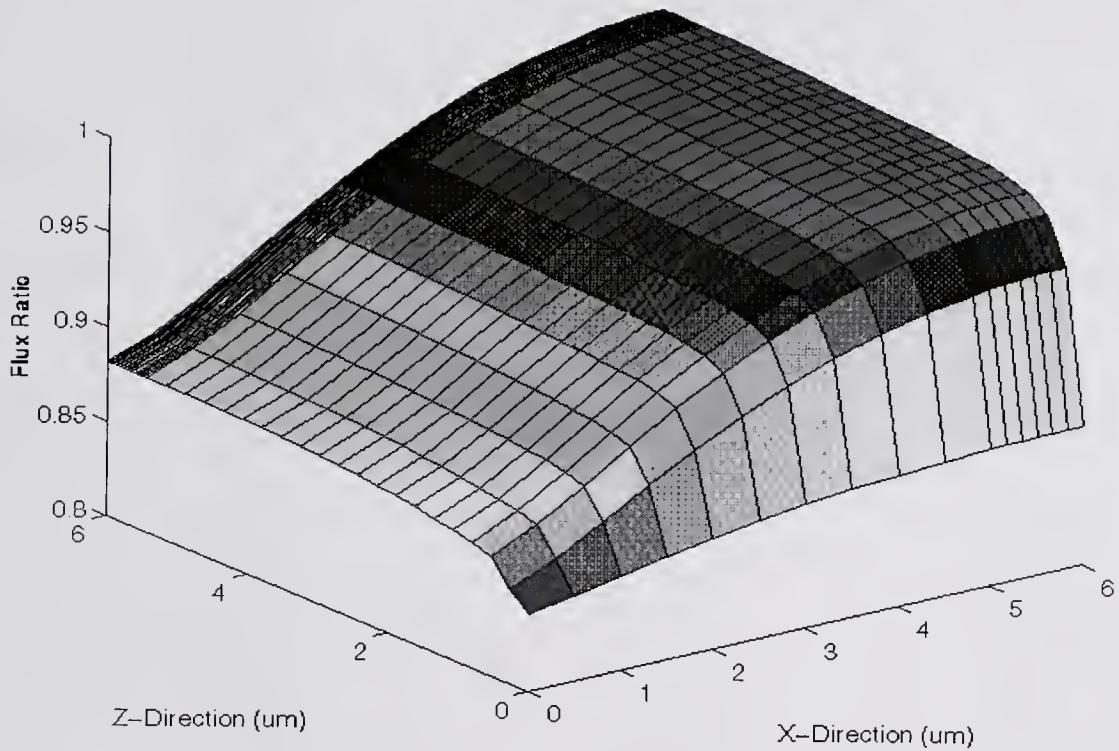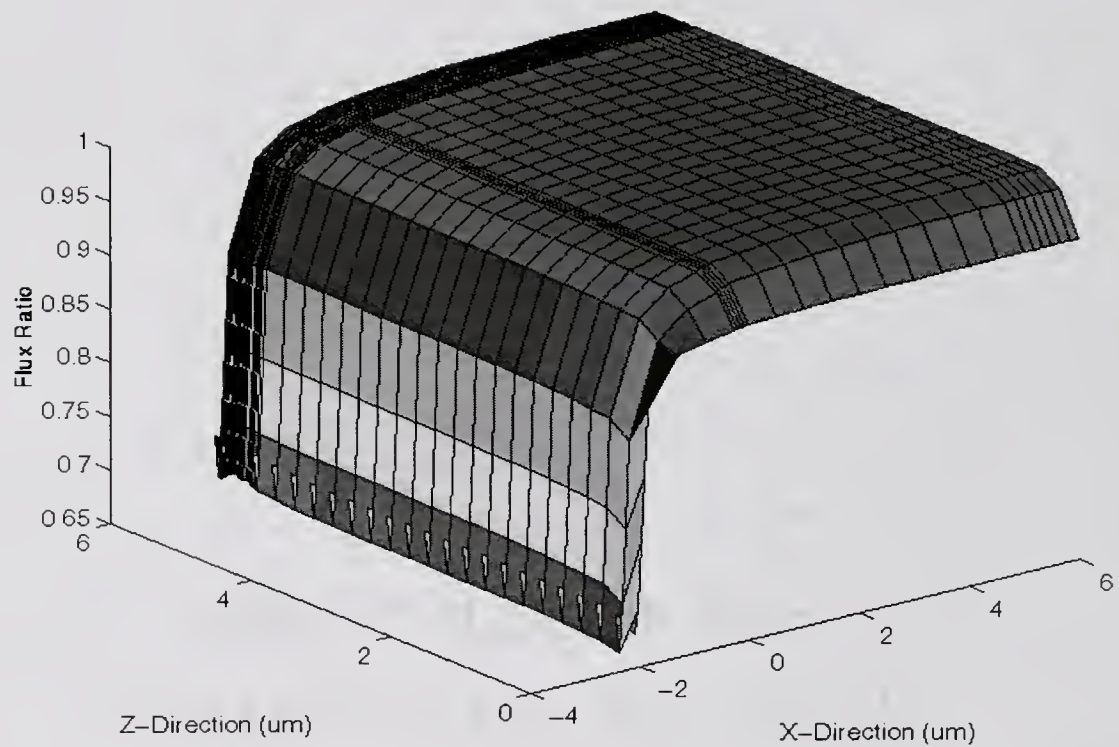
Figure 4.5   The ratio of the y-component of the heat flux vector to the magnitude of the total heat flux vector in the exterior silicon at the trench/exterior silicon interface. The device parameters are $L = 2 \, \mu m$, $C_L = 10 \, \mu m$, $d_{trox} = 0.25 \, \mu m$ and $d_{poly} = 1.6 \, \mu m$.

The constant C_n is $(L_{tub} + 2d_{tr})/(2 \tan \theta_L)$ for the x-direction and $(W_{tub} + 2d_{tr})/(2 \tan \theta_W)$ for the y-direction, where θ_W and θ_L are the thermal spreading angles for the cooling fins which are assumed to be 45° [Hir93]. The third term on the left-hand side of (4.45) accounts for heat lost from the exterior silicon due to conduction through the buried oxide, where

$$\frac{1}{m_{si}} = \sqrt{\frac{d_{tub}}{H_{ox}(t)}} \tag{4.46}$$

is the characteristic thermal length in the exterior silicon. Equation (4.45) does not have a simple closed-form solution. However, a solution to the steady-state form of (4.45) exists and, by using the time-dependent characteristic thermal length, can be used to approximate the temperature rise in the exterior silicon as

$$\Delta T_{si} = c_1 K_0[m_{si}(t) \cdot (n + C_n)], \tag{4.47}$$

where c_1 is an arbitrary constant and K_i is the modified Bessel function of the second kind of order i.

The equation that describes the flux at the interface between the tub and the trench is derived by solving (4.40) using (4.31), (4.41) through (4.44), and (4.47). The resulting solution, when arranged into the form given by (4.4), yields the normalized heat-transfer coefficient for the trench

$$H_{tr}(t) = \frac{k_{tr}m_{si}K_1(m_{si}C_n)}{\left[k_{si}m_{si}\sqrt{\pi\alpha_{tr}t} \cdot erf\left(\frac{d_{tr}}{\sqrt{4\alpha_{tr}t}}\right)K_1(m_{si}C_n) + k_{tr}exp\left(\frac{d_{tr}^2}{4\alpha_{tr}t}\right)K_0(m_{si}C_n)\right]} \quad (4.48)$$

4.2.3 Effects of Interconnect Metallization on the Thermal Impedance

For the derivation of the DIBJT thermal impedance model, the surface of the device is assumed to be adiabatic, and therefore conduction through the interconnect metallization is neglected. Since the metallization typically has a high thermal conductivity, it is possible that the heat conduction via the interconnects significantly influences the thermal impedance. Therefore, the validity of such an assumption should be investigated.

Based on the finite-element (FE) simulations performed for bulk BJT's (Chapter Two), the heat conduction through base and collector interconnects is typically negligible and only the emitter interconnect is assumed to affect the thermal response of the device. To determine the extent of the effect of the emitter interconnect on the thermal impedance, both steady-state and transient three-dimensional (3-D) FE simulations of a DIBJT structure were performed using ANSYS. The FE model represented a simplified device structure and was constructed using similar assumptions and boundary conditions to those detailed in Chapter Two. Figure 4.6a shows the FE model used for the thermal simulations. The simulations were carried out with and without the emitter interconnect present, while independently varying each structure parameter of the FE device model.
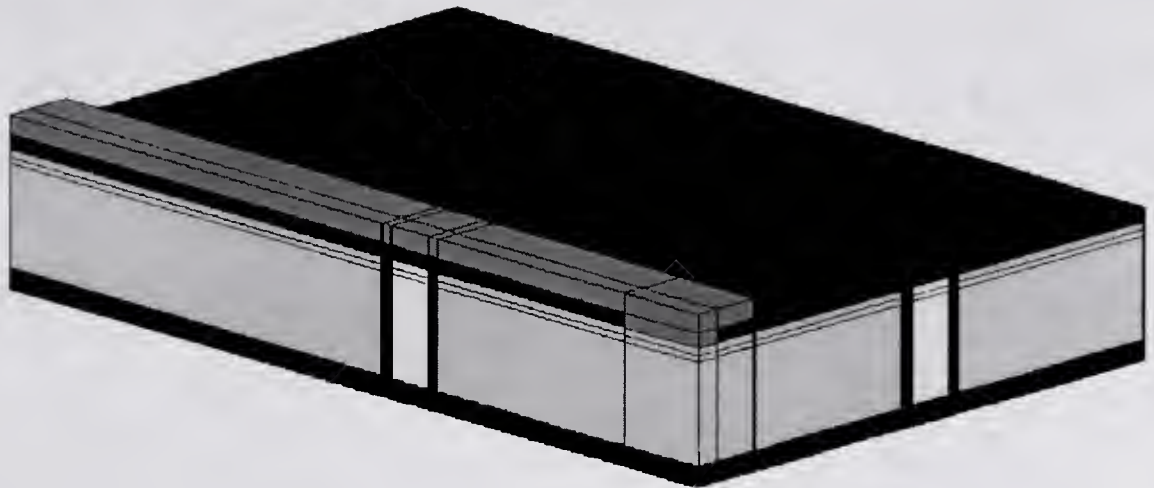
(a)
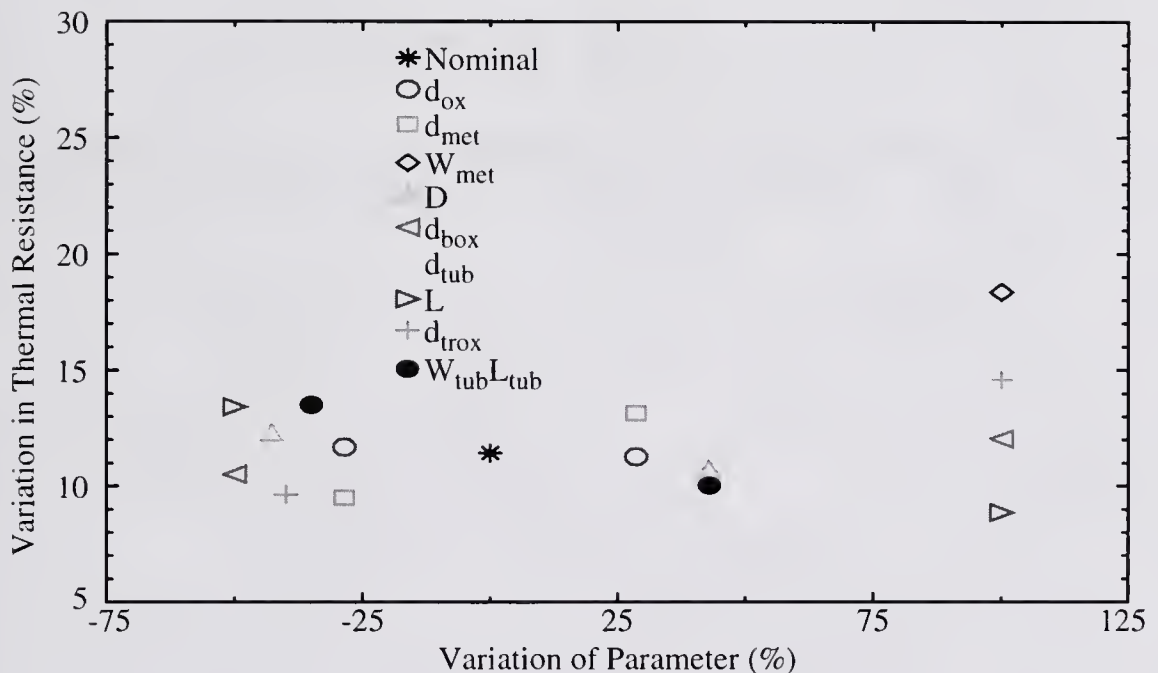
(b)

Figure 4.6   ANSYS simulations showing the effect of the emitter interconnect on the thermal resistance. The specifications, in microns, for the nominal device are $L = 4$, $W = 1$, $D = 0.35$, $H = 0.3$, $d_{box} = 1$, $d_{tub} = 5.5$, $L_{tub} = 12$, $W_{tub} = 9$, $d_{trox} = 0.25$, $d_{poly} = 1.6$, $d_{ox} = 0.7$, $d_{met} = 0.7$, $W_{met} = 2$: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for the emitter interconnect and the thermal resistance neglecting the emitter interconnect. The variation plotted on the x-axis corresponds to the deviation of each structure parameter from its nominal value.

The results of the steady-state FE simulations are shown in Figure 4.6b. As with the bulk BJT's simulated in Chapter Two, the emitter interconnect enhances the heat flow away from the active device area, resulting in lower thermal resistance values. As was the case with the bulk BJT FE model simulated in Chapter Two, the effectiveness of the interconnect as a thermal conductance path increases as its cross-sectional area increases or as the interconnect is effectively moved closer to the heat source. However, the influence of the emitter interconnect on the thermal resistance is more significant for the DIBJT's than for the bulk BJT's. This effect is due to the lower thermal conductance to the substrate of DIBJT's as compared to the bulk devices. Therefore, the conductance through the emitter interconnect becomes a larger component of the total thermal conductance of the device. This trend is illustrated by the data in Figure 4.6b, which shows that as the conductance to the substrate is increased by enlarging the tub or thinning the isolation oxides, the effect of the emitter interconnect on the thermal resistance is reduced.

Since the effect of the emitter interconnect is increased due to the dielectric isolation, it is reasonable to suggest that the effects of the base and collector interconnects on the thermal resistance will also increase. Therefore, the assumption that the base and collector metallization is negligible becomes questionable. However, the increase in the effects of the base and collector interconnects should not be any greater than that for the emitter interconnect. Consequently, even for a two-fold increase in the influence of the base and collector metallization in a bulk technology, the heat conduction through the base and collector interconnects will still be negligible.

Transient thermal simulations were performed to examine the effect of the emitter interconnect on the dynamic temperature response. ANSYS was used to simulate the structure in Figure 4.6a for a step increase in power dissipation, both with and without the emitter interconnect in contact with the device. Figure 4.7 shows the results of the transient FE simulations. The observed trend in the transient temperature rise is similar to that for the bulk bipolar devices in Chapter Two. The additional heat capacity of the emitter metallization effectively slows the temperature rise in the device. The extent of the metallization's effect on the dynamic temperature response is directly related to the effective volume of the interconnect structure, and will be more pronounced for large devices with substantial contact area. Therefore, the model that neglects the conduction through the interconnect by assuming an adiabatic surface will produce the quickest possible temperature rise.

4.2.4 A Model for the Thermal Impedance of the Emitter Interconnect

As detailed in Chapter Two, the effect of the emitter interconnect on the thermal response can be modeled by an additional thermal impedance. This thermal impedance is derived by calculating both the thermal resistance and thermal capacitance of the emitter metallization.

The emitter interconnect is represented by a one-dimensional cooling fin, such that the thermal resistance can be expressed as

$$R_{THmet} = [k_{met}m_{met}W_{met}d_{met}]^{-1}$$
(4.49)
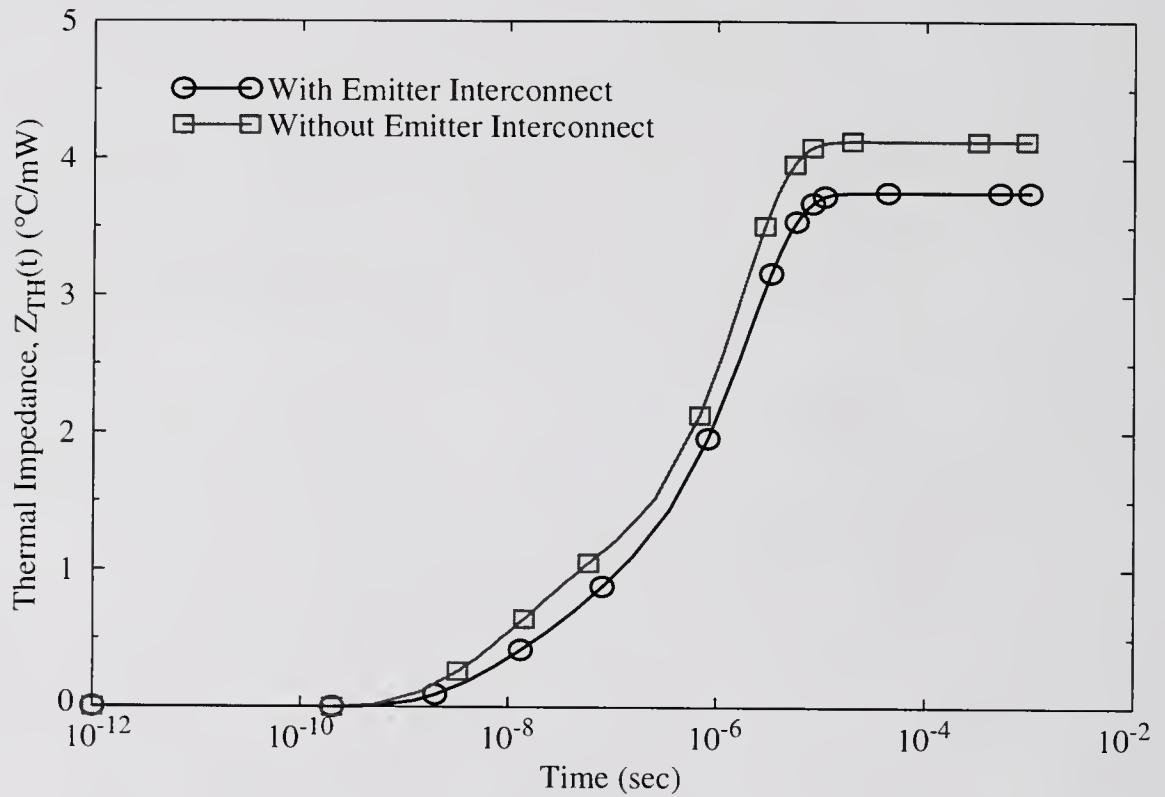
Figure 4.7   ANSYS simulations showing the effect of the emitter interconnect on the thermal impedance. The simulated FE device model used the nominal structure parameters listed in the caption of Figure 4.6.

where

$$\frac{1}{m_{met}} = \sqrt{\frac{k_{met}d_{met}}{h_{met}}} \qquad (4.50)$$

is the characteristic thermal length in the interconnect and

$$h_{met} = \frac{k_{ox}k_{si}}{k_{si}(d_{ox} + d_{box}) + k_{ox}d_{tub}} \qquad (4.51)$$

is the heat-transfer coefficient from the interconnect to the substrate. The material properties for the emitter interconnect are given in Table 4.3.

Table 4.3 Emitter interconnect[‡] material properties

Property	Definition	Value
k_{met}	Thermal conductivity	2.39 (W cm^{-1} K^{-1})
ρ_{met}	Density	2.7 (g cm^{-3})
c_{pmet}	Specific heat	0.9 (J g^{-1} K^{-1})

Source: [Ozi93]
[‡] Assumed to be aluminum

The thermal capacitance of the emitter metallization is given by

$$C_{THmet} = \rho_{met}c_{pmet}V_{met} \qquad (4.52)$$

where $V_{met} = W \cdot L \cdot \delta_{met}$ is the effective volume of the interconnect structure.

Therefore, the transient thermal impedance of the emitter interconnect can be approximated as

$$Z_{THmet}(t) = R_{THmet}\left[1 - \exp\left(\frac{-t}{\tau_{met}}\right)\right] \qquad (4.53)$$

where $\tau_{met} = R_{THmet}C_{THmet}$. The overall thermal impedance of a bipolar device can now be represented by the parallel combination of two thermal impedances, such that effectively

$$Z_{TH}(s) = \frac{Z_{THdev}(s) \cdot Z_{THmet}(s)}{Z_{THdev}(s) + Z_{THmet}(s)}, \qquad (4.54)$$

where $Z_{THdev}(s)$ is determined from the transient thermal impedance given by (4.23).

4.3 Verification of the DIBJT Thermal Impedance Model

To verify the thermal impedance model, three-dimensional (3-D) finite-element (FE) simulations of a DIBJT were performed using ANSYS. The FE model represented a simplified device structure and was constructed using similar assumptions and boundary conditions to those detailed in Chapter Two. In addition, all interconnect metallization was neglected and the tub region was assumed to be composed of homogeneous silicon with the bulk properties given in Table 4.2. The FE simulations were evaluated at the surface corner of the emitter and compared to (4.23); the results are shown in Figure 4.8. The DIBJT thermal impedance model
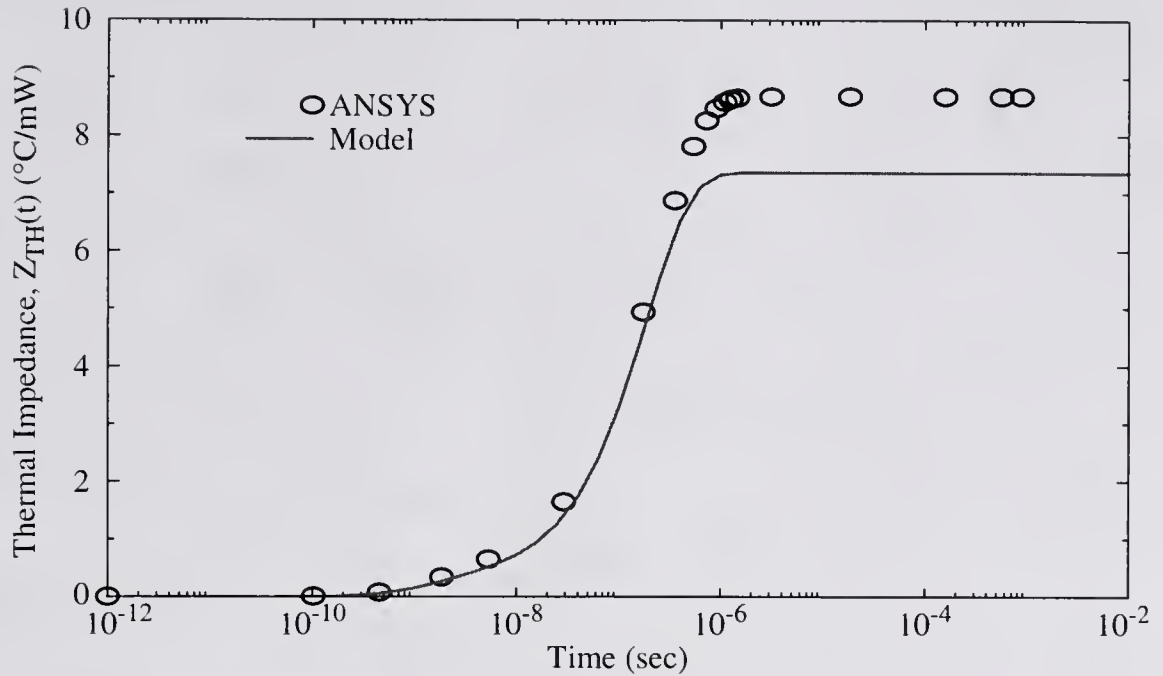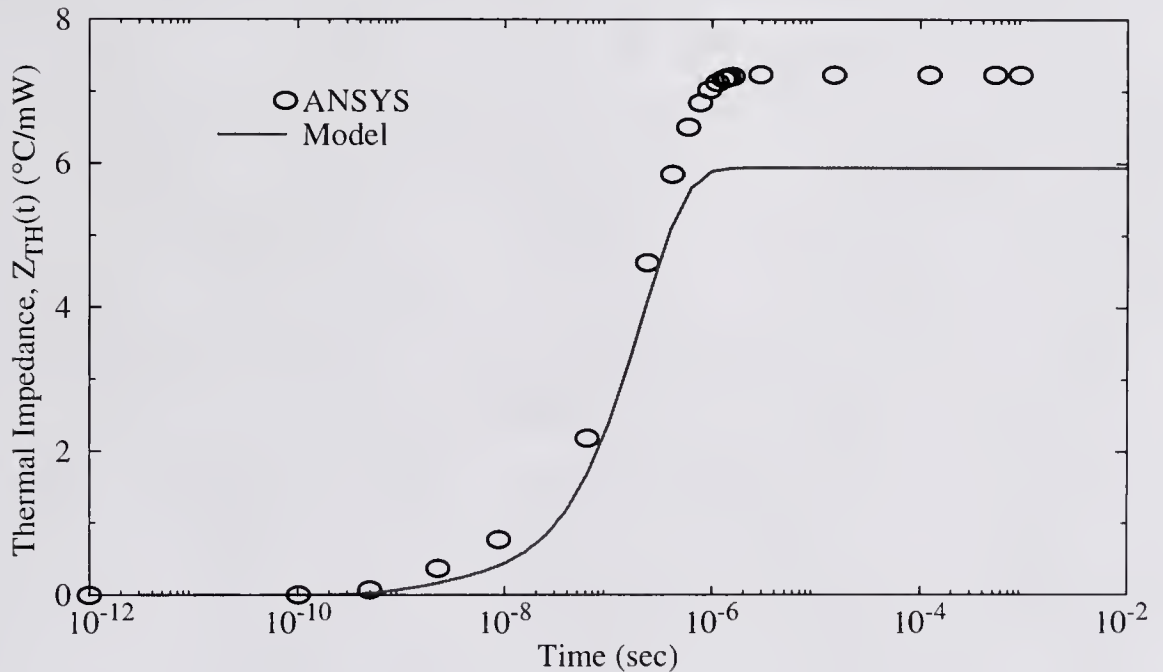
Figure 4.8   The transient thermal impedance simulated with ANSYS and calculated with the DIBJT model for d_{tub} = 1.5 μm, d_{box} = 0.5 μm, d_{trox} = 0.13 μm and d_{poly} = 0.5 μm: a) L = 2 μm and W = 0.5 μm; b) L = 3 μm and W = 0.7 μm

agrees closely, for both the transient and steady-state, with the 3-D FE simulation results. The largest error, which is approximately 17%, occurs in steady-state and can be partially attributed to numerical error associated with limitations of the FE mesh.

The model was also compared to measured thermal impedance data extracted by Zweidinger et al. using a base-current thermometry technique [Zwe96]. Figure 4.9, Figure 4.10 and Figure 4.11 compare the measured and simulated data for the transient thermal impedances of Harris UHF DIBJT's. The thermal impedance model does a good job of predicting the steady-state thermal resistance, with no more than 11% error between the model and the measurements. However, the model given by (4.23) tends to exaggerate the transient response. This discrepancy can be attributed to the model's neglect of the emitter interconnect metallization. Equation (4.23) predicts the quickest temperature response in the device since it does not account for the additional heat capacity of the metallization. However, when the thermal impedance of the emitter interconnect is accounted for, with $\delta_{met} = 100 \, \mu m$ extracted from the measurements, the model provides a more accurate representation of the transient temperature response.

4.4 Derivation of a Compact DIBJT Thermal Resistance Model

For the limiting case of steady-state thermal conduction, the complexity of the DIBJT thermal impedance model can be reduced, resulting in a single closed-form expression for the thermal resistance. The approach for the derivation of the thermal resistance model is adopted from the analysis by Goodson and Flik for SOI
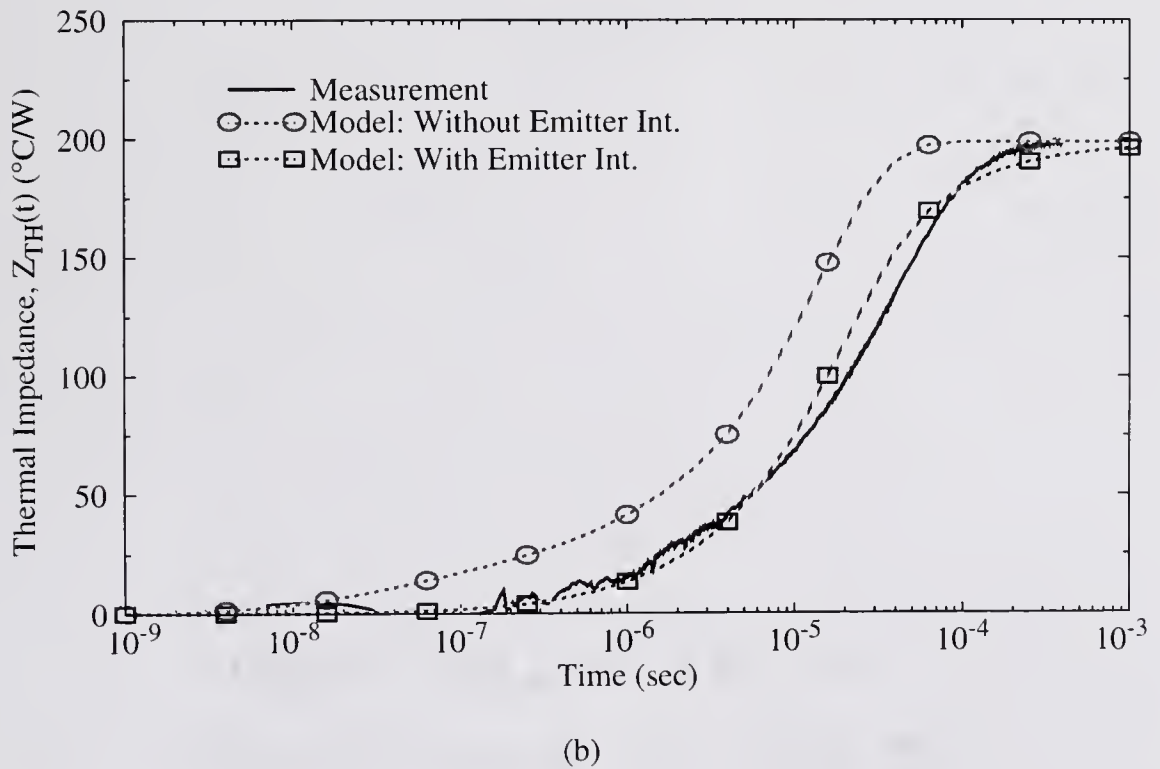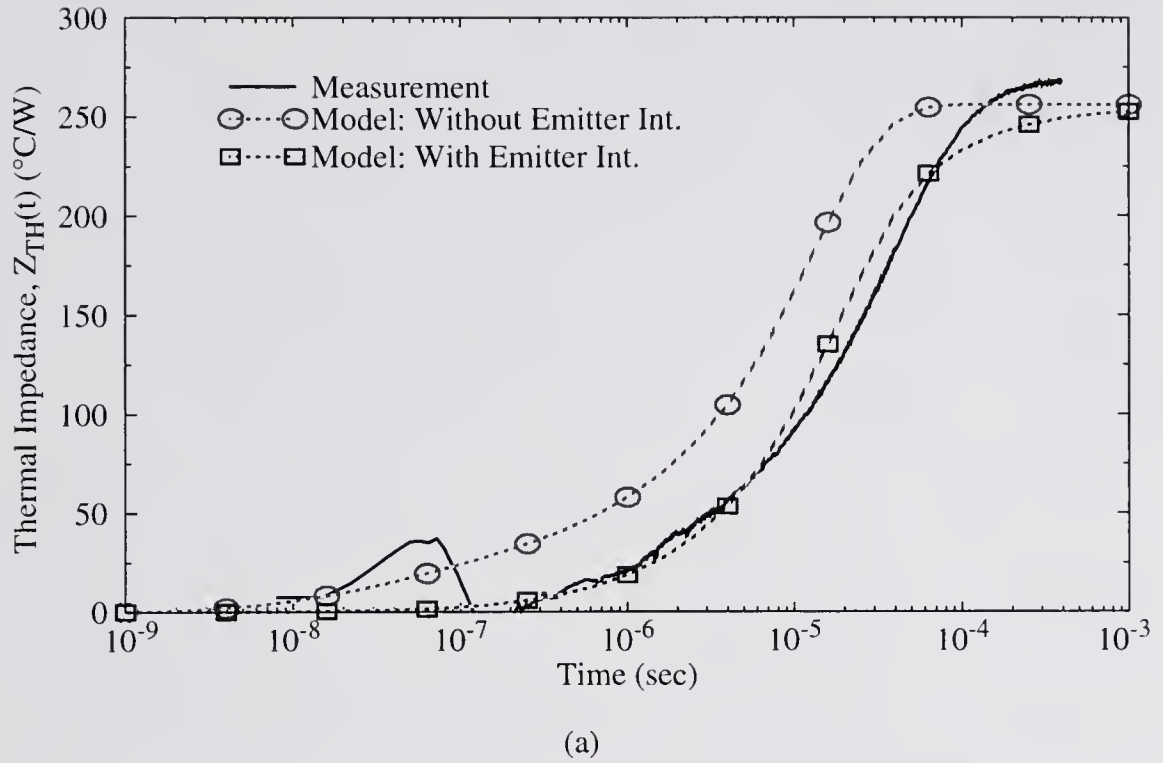
Figure 4.9   Measured and simulated data for the transient thermal impedance of Harris UHF DIBJT's with W = 3 μm: a) L = 50 μm; b) L = 70 μm.
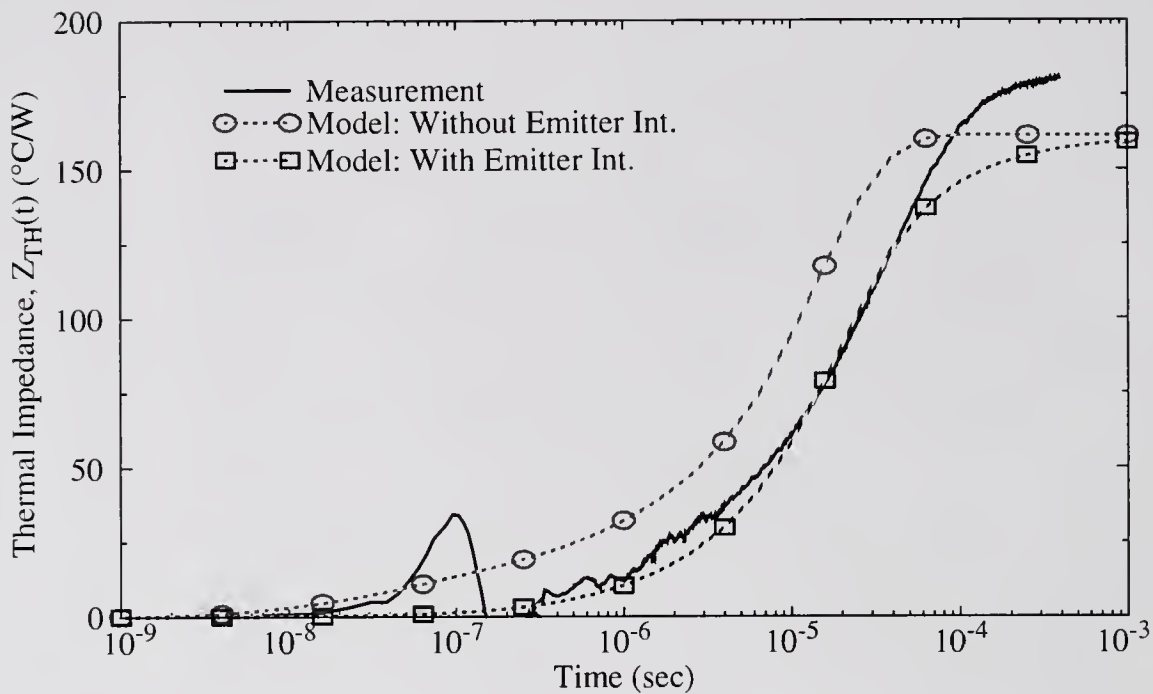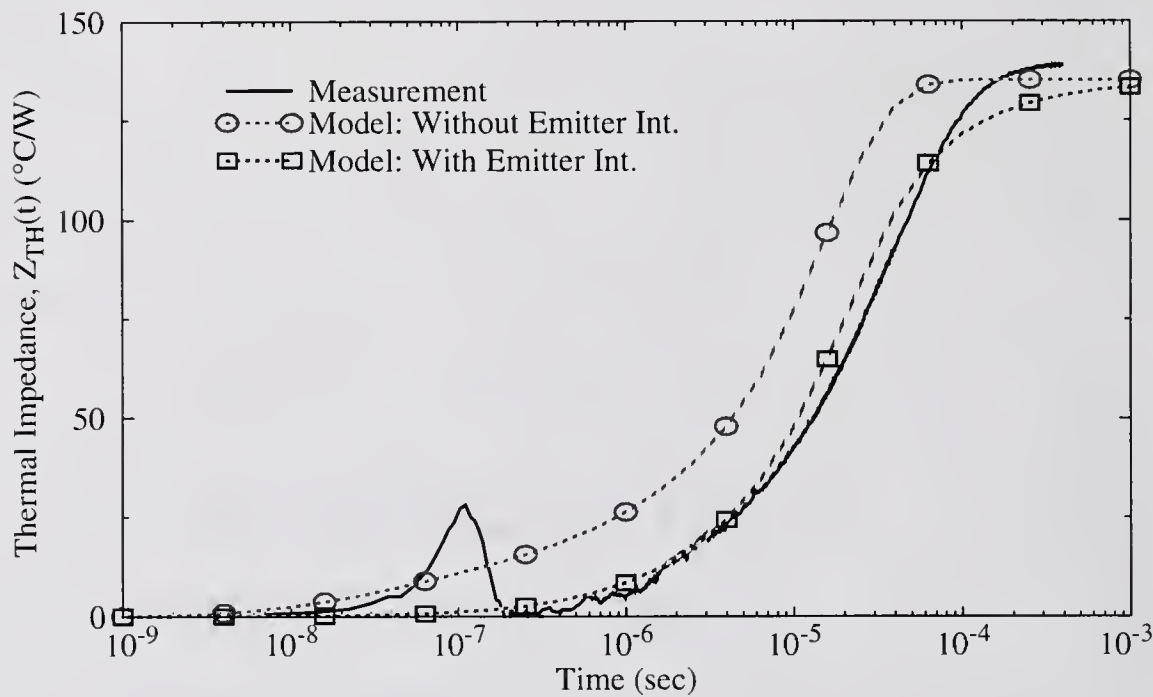
Figure 4.10   Measured and simulated data for the transient thermal impedance of Harris UHF DIBJT's with W = 3 μm: a) L = 90 μm; b) L = 110 μm.
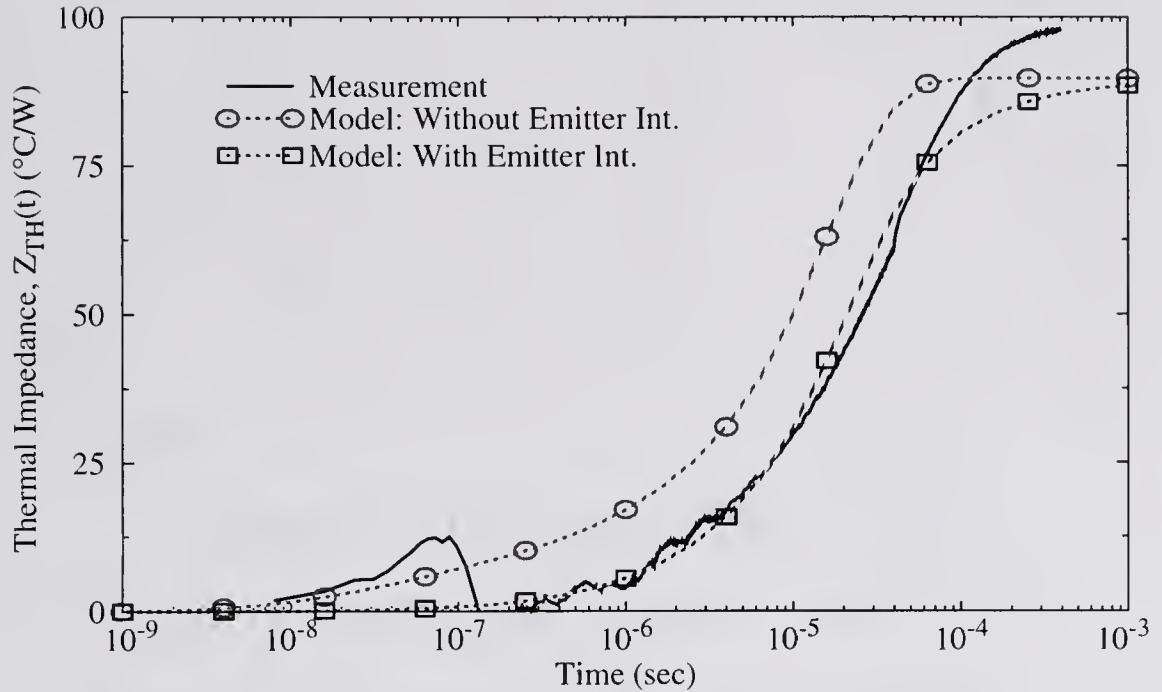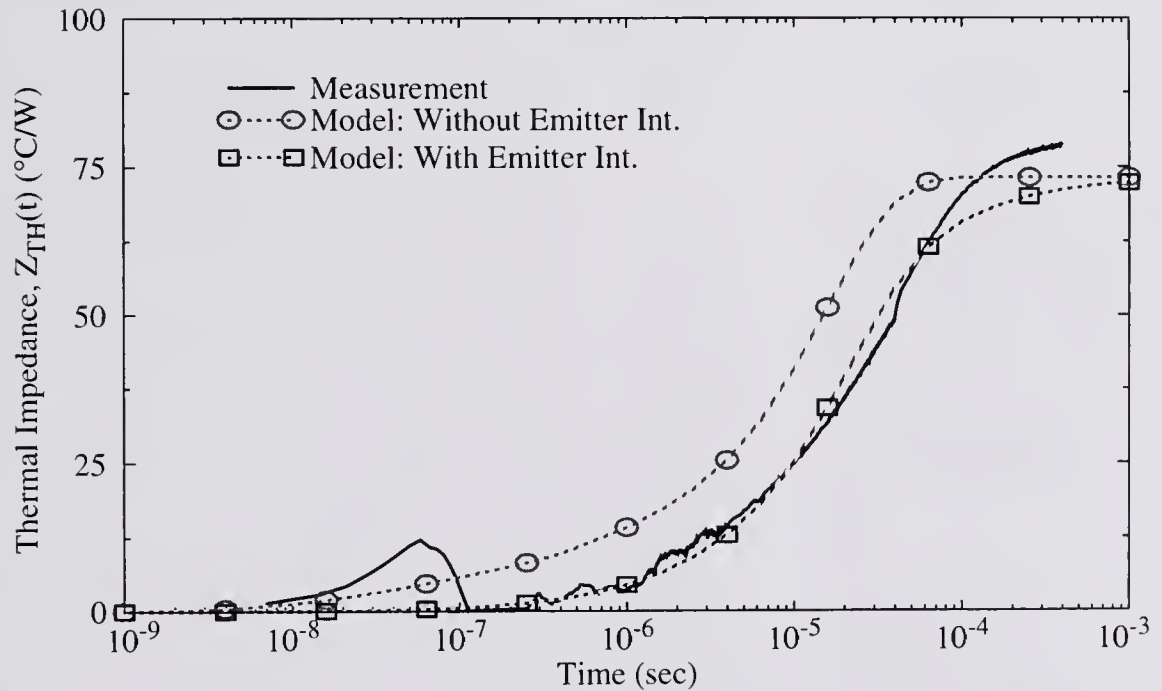
Figure 4.11   Measured and simulated data for the transient thermal impedance of Harris UHF DIBJT's with W = 3 μm: a) L = 170 μm; b) L = 210 μm.

MOSFET's [Goo92]. The model represents a solitary DIBJT device with three parallel conductance paths that carry heat away from the tub. These heat-flow paths are illustrated in Figure 4.12 and are modeled by a system of coupled one-dimensional differential equations. The derivation of the thermal resistance model is divided into separate regions that correspond to the different thermal conductance paths. Different subscripts are used to denote the equations and variables that apply to the different regions; the subscripts tub, e and si refer to the tub region, emitter interconnect and exterior silicon region, respectively. Figure 4.13 shows the simplified device structure that defines the geometry for the model derivation. The definitions of the model parameters are listed in Table 4.4.

The defining assumption for the DIBJT thermal resistance model, and the main departure from the thermal impedance model, is that the entire tub is considered to be a heat source, with uniform power dissipation equal to the actual device power P, at a single uniform temperature, T_{tub}. Since the thermal conductivity of the tub silicon is much greater than that of the insulating oxide, the thermal conductance of the tub is larger than that of the trenches or the buried oxide, and the thermal resistance of the device is mainly determined by the isolation structures. This assumption is supported by both ANSYS simulations and the MEDICI simulation in the work by Ganci et al. [Gan92]. Figure 4.14 shows the ANSYS results, which illustrate that the temperature gradient in the tub is smaller than the gradients across the isolation structures. However, the validity of this assumption declines as the dimensions of the tub, $A_{tub} = W_{tub}L_{tub}$, become much greater than those of the active device, $A_{dev} = W \cdot L$, or the thickness of the insulating oxide layers (buried
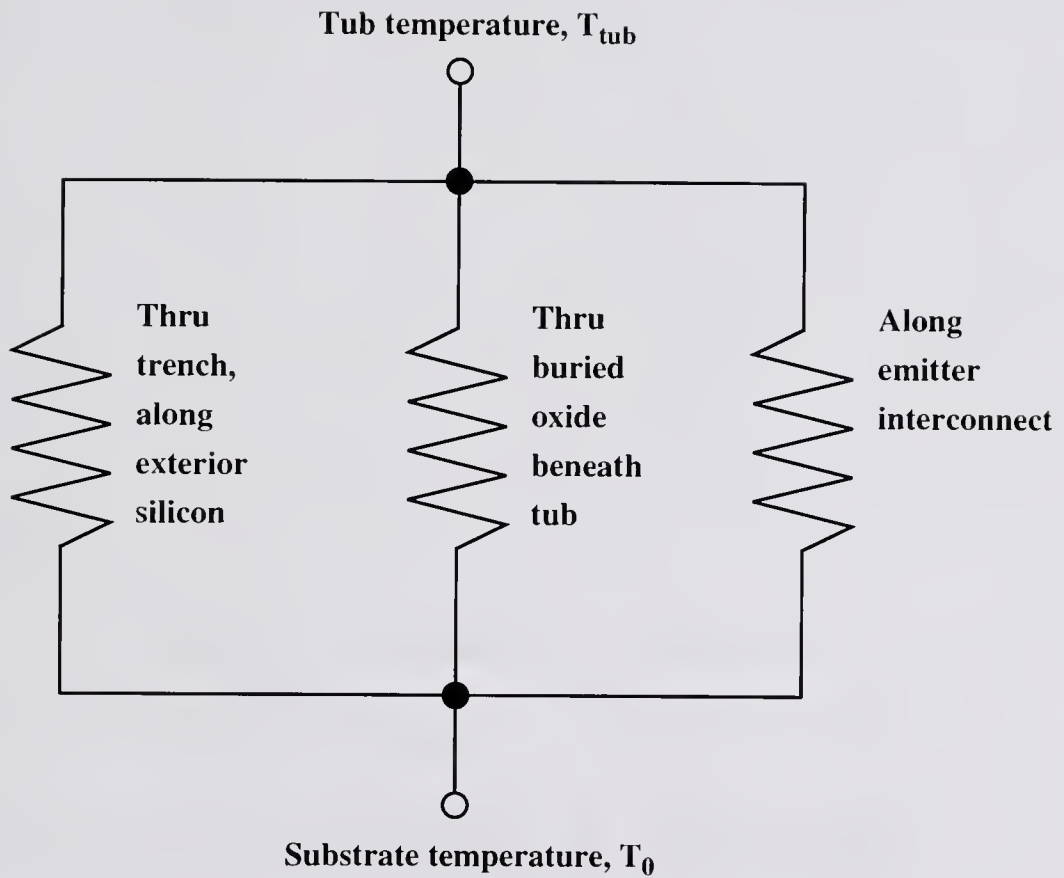
Figure 4.12   Illustration of the three parallel conductance paths for heat to travel from the device tub to the semiconductor substrate.
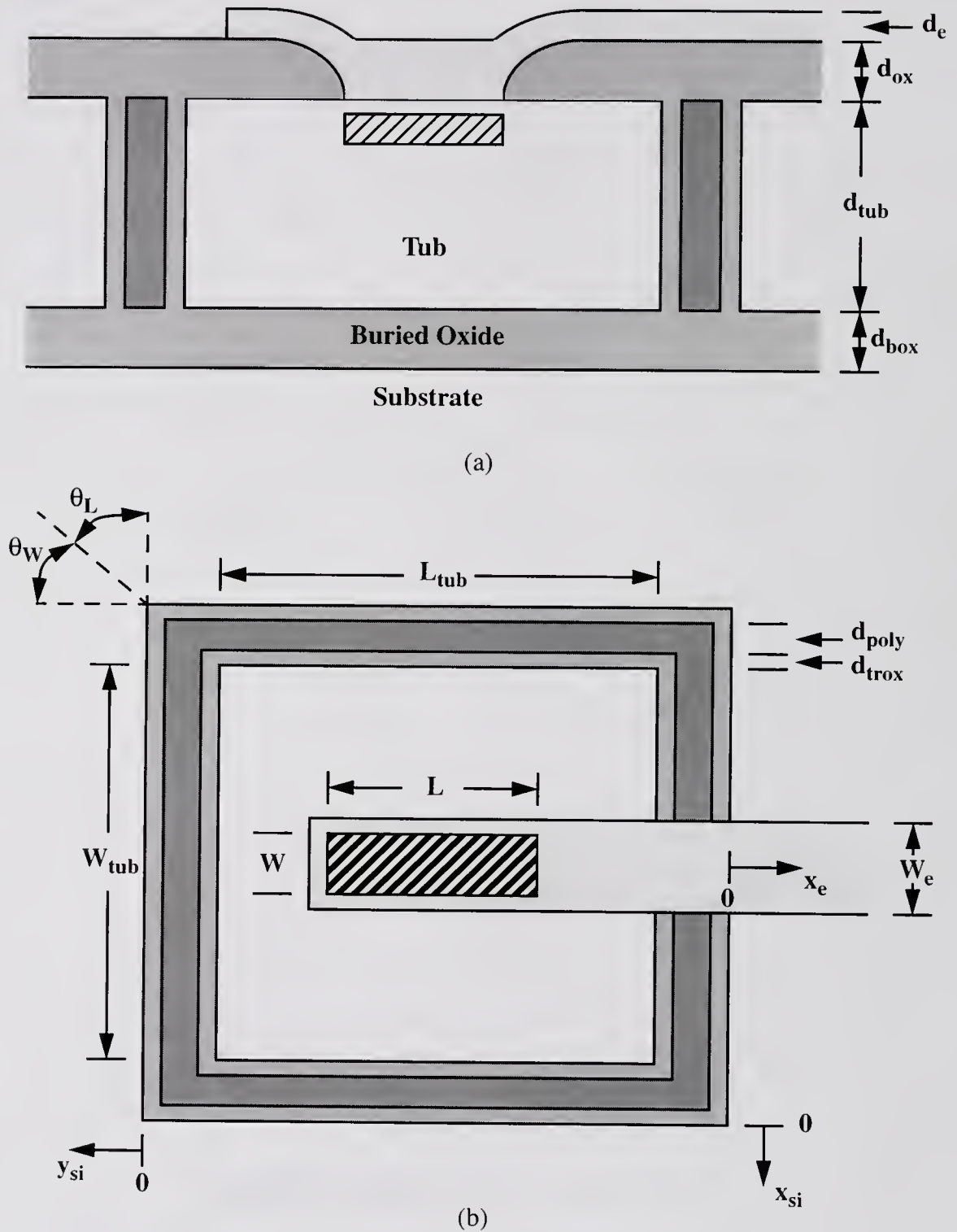
Figure 4.13   The simplified device geometry used to define the solution domain for the DIBJT thermal resistance model. The model parameters are defined in Table 4.4: a) Cross-sectional view; b) top view.

Table 4.4 DIBJT thermal resistance model parameters

Parameter	Definition
W	Width of emitter stripe
L	Length of emitter stripe
d_{poly}	Thickness of trench polysilicon
d_{trox}	Thickness of trench oxide
d_{tub}	Thickness of silicon tub
d_{box}	Thickness of buried oxide
d_{ox}	Total thickness of oxide between emitter interconnect and device surface
d_e	Thickness of emitter interconnect
W_{tub}	Width of silicon tub
L_{tub}	Length of silicon tub
W_e	Width of emitter interconnect
C_W	Scaling constant for tub width
C_L	Scaling constant for tub length
C_e	Scaling constant for emitter interconnect width
θ_W	Spreading angle of heat flow
θ_L	Spreading angle of heat flow

(a)

(b)
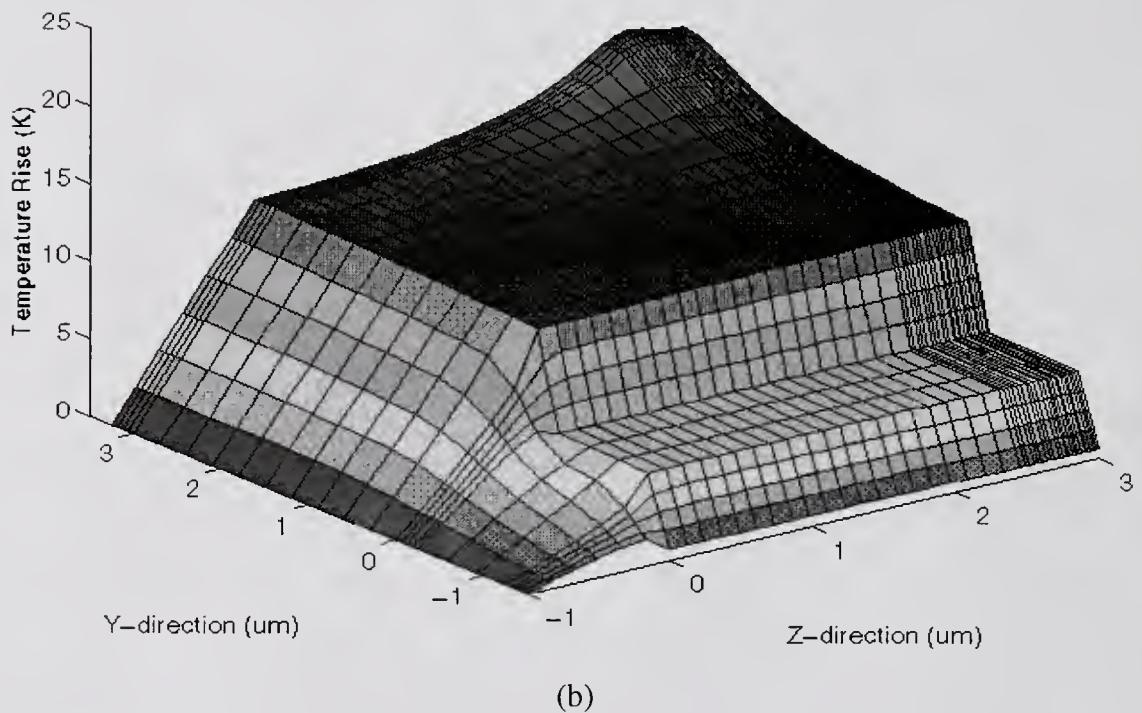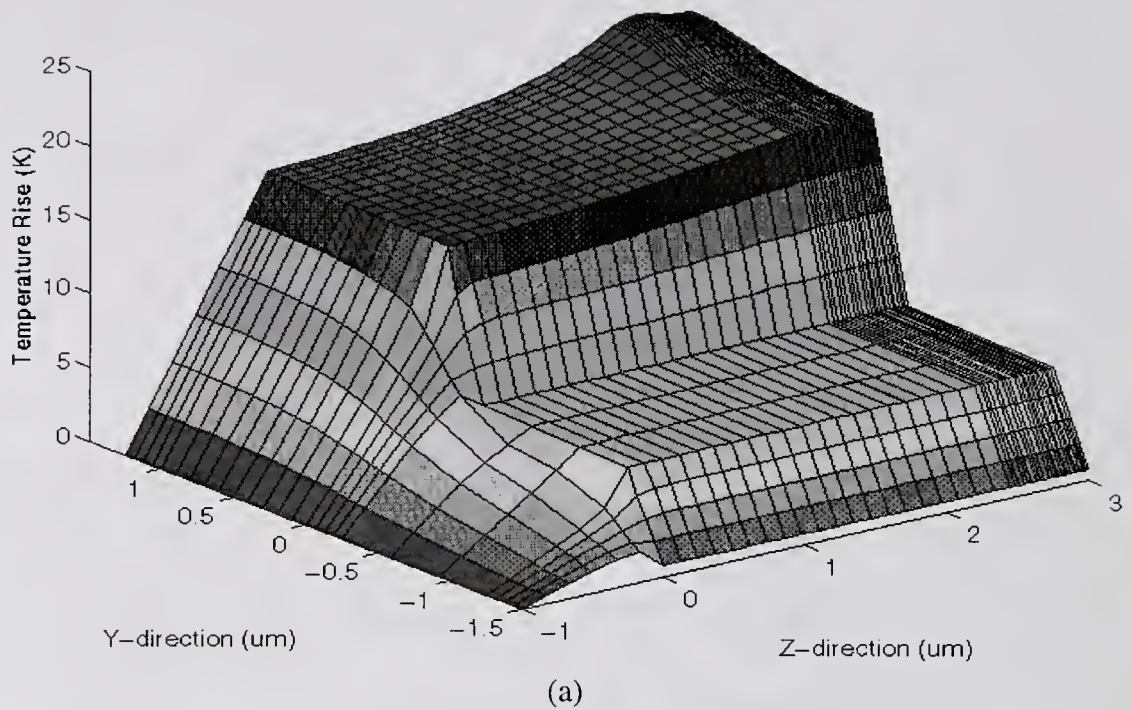
Figure 4.14   ANSYS simulation results showing the temperature gradient in the tub and across the isolation structures: a) For $A_{tub}/A_{dev} = 9$; b) for $A_{tub}/A_{dev} = 18$.

or trench) is reduced. For such conditions, the conductance of the tub is reduced and the temperature gradients within the tub become more significant, so that the model will tend to underestimate the thermal resistance.

The boundary conditions and the model parameters for the thermal resistance model, for the most part, are the same as those for the thermal impedance model. The interface between the buried oxide and the substrate is assumed to be at a uniform temperature T_0, and the device is assumed to cool solely through the substrate. The width and length of the tub region are assumed to scale directly with the width and length of the emitter stripe; one change from the thermal impedance model is that the scaling constants C_{W1} and C_{W2} have been lumped together so that

$$W_{tub} = W + C_W. \tag{4.55}$$

The top surface of the device is considered adiabatic so that heat conduction through the overlying oxide layers is neglected.

Based on the finite-element simulations in Chapter Two, only the emitter interconnect is assumed to affect the thermal resistance. However, since the tub is assumed to be at a uniform temperature, the effects of the base and collector metallization can be incorporated using a similar analysis as shown below. The emitter interconnect is still treated as a one-dimensional cooling fin. Therefore, the temperature rise along the interconnect can be described by the following differential equation

$$\frac{\partial^2 \Delta T_e}{\partial x_e^2} - m_e^2 \Delta T_e = 0 \tag{4.56}$$

and the boundary conditions

$$\Delta T_e \big|_{x_e = 0} = \Delta T_{tub} \tag{4.57}$$

$$\Delta T_e \big|_{x_e \to \infty} = 0, \tag{4.58}$$

where $\Delta T_e(x_e) = T_e(x_e) - T_0$. The assumption that the temperature gradient in the emitter interconnect is primarily one-dimensional was validated in Chapter Two. The width of the emitter interconnect is assumed to scale directly with the width of the emitter stripe

$$W_e = W + C_e, \tag{4.59}$$

where C_e depends on the fabrication process. The characteristic thermal length and the heat-transfer coefficient for the interconnect are given by (4.50) and (4.51), respectively.

The heat lost from the tub through the trench is governed by the steady-state heat-transfer coefficient of the trench. The temperature gradient in the trench is assumed to be one-dimensional and the exterior silicon is modeled by four one-dimensional cooling fins where for steady-state, the characteristic thermal length is

$$\frac{1}{m_{si}} = \sqrt{\frac{k_{si} d_{tub}}{h_{ox}}} \tag{4.60}$$

and the buried-oxide heat-transfer coefficient is given by

$$h_{ox} = \frac{k_{ox}}{d_{box}}. \tag{4.61}$$

The flux at the interface between the tub and the trench can be expressed as

$$-k_{tr}\frac{\partial \Delta T_{tr}}{\partial n} = h_{tr} \cdot \Delta T_{tub}. \tag{4.62}$$

Equation (4.62) can be solved for the heat-transfer coefficient of the trench, h_{tr}, using the steady-state form of (4.40), equations (4.41) through (4.44) and (4.47), resulting in

$$h_{tr} = \frac{k_{tr}k_{si}m_{si}K_1(m_{si}C_n)}{d_{tr}k_{si}m_{si}K_1(m_{si}C_n) + k_{tr}K_0(m_{si}C_n)}. \tag{4.63}$$

The thermal conductance paths are coupled through the following power conservation equation

$$-k_e W_e d_e \frac{\partial \Delta T_e}{\partial x_e}\bigg|_{x_e = 0} + h_{ox}W_{tub}L_{tub}\Delta T_{tub} + 2h_{tr}d_{tub}(W_{tub} + L_{tub})\Delta T_{tub} = P. \tag{4.64}$$

Equation (4.64) can be solved in the form

$$\Delta T_{tub} = R_{TH}P, \tag{4.65}$$

resulting in the following expression for the thermal resistance

$$R_{TH} = [k_e W_e d_e m_e + h_{ox} W_{tub} L_{tub} + 2h_{tr} d_{tub}(W_{tub}\gamma_y + L_{tub}\gamma_x)]^{-1} \qquad (4.66)$$

$$\gamma_y = \frac{k_{si}m_{si}K_1\left[\dfrac{m_{si}(W_{tub} + 2d_{tr})}{2\tan\theta_W}\right]}{h_{tr}K_0\left[\dfrac{m_{si}(W_{tub} + 2d_{tr})}{2\tan\theta_W}\right] + k_{si}m_{si}K_1\left[\dfrac{m_{si}(W_{tub} + 2d_{tr})}{2\tan\theta_W}\right]} \qquad (4.67)$$

$$\gamma_x = \frac{k_{si}m_{si}K_1\left[\dfrac{m_{si}(L_{tub} + 2d_{tr})}{2\tan\theta_L}\right]}{h_{tr}K_0\left[\dfrac{m_{si}(L_{tub} + 2d_{tr})}{2\tan\theta_L}\right] + k_{si}m_{si}K_1\left[\dfrac{m_{si}(L_{tub} + 2d_{tr})}{2\tan\theta_L}\right]} \cdot \qquad (4.68)$$

The first term on the right-hand side of (4.66) corresponds to the heat flow out through the emitter interconnect; the second term corresponds to the heat flow out through the buried oxide, and the third term corresponds to the heat flow out through the trench.

4.5 Verification of the DIBJT Thermal Resistance Model

The accuracy of the thermal resistance model is tested with comparisons to measured steady-state thermal resistances and to the thermal impedance model; the results are shown in Figure 4.15. The thermal resistance model displays the proper trends with emitter length and tub scaling but as A_{tub}/A_{dev} increases, the accuracy of the model declines. For the DIBJT's with $A_{tub}/A_{dev} > 30$, the error between the thermal resistance model and the measured data is in excess of 20%. This trend is further illustrated by examining the thermal resistance of Harris
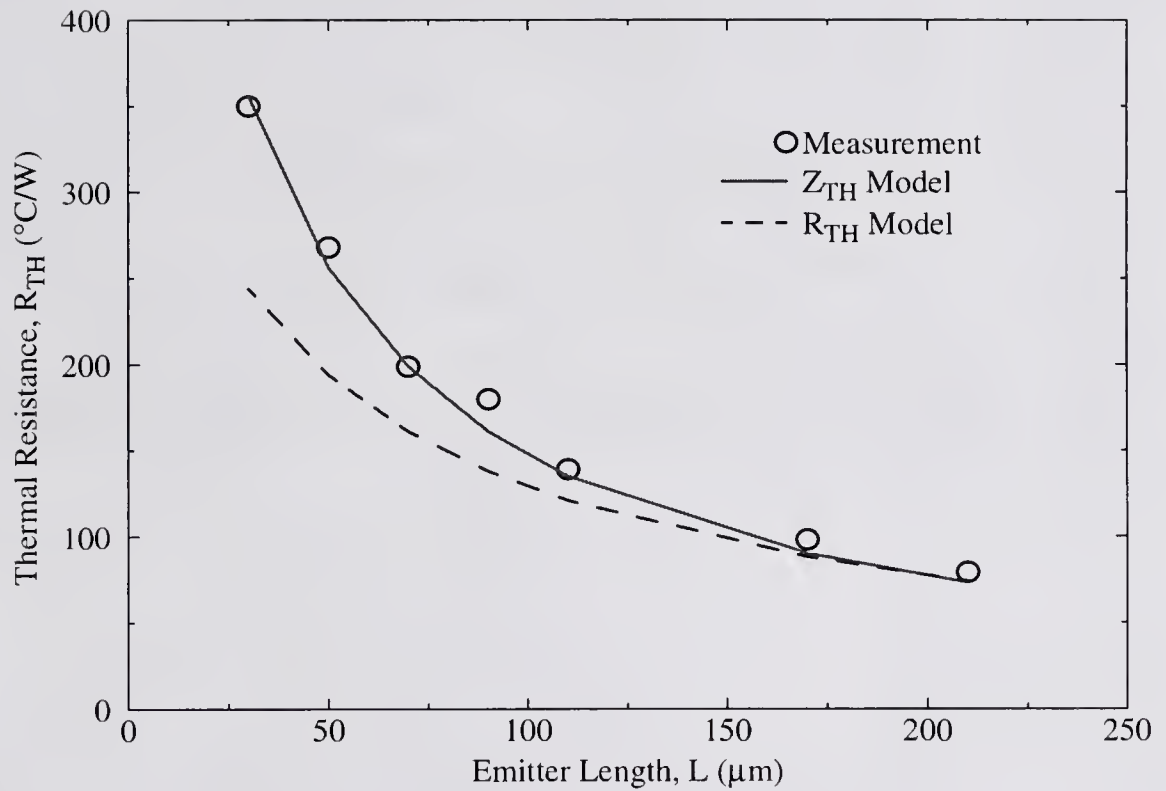
Figure 4.15   The steady-state thermal resistance extracted from measurements and calculated using both the thermal impedance and thermal resistance models; W = 3 μm.

"Cooling-Zone" DIBJT's. These devices were fabricated in larger tubs to reduce the thermal resistance. Table 4.5 shows a comparison of the measured and predicted thermal resistances for two "Cooling-Zone" devices.

Table 4.5 Thermal resistance (°C/W) of Harris "Cooling-Zone" DIBJT's[‡]

L (μm)	A_{tub}/A_{dev}	Measured	Z_{TH} Model	R_{TH} Model
62	78	150	165	82
125	118	93	80	30

[‡] $W = 3 \ \mu m$

As predicted, the thermal resistance model significantly under-estimates the measured thermal resistance since the tub area is much greater than the device area. The thermal impedance model, which accounts for the temperature gradients in the tub, provides a better estimate for the thermal resistance. In the work by Ganci et al. [Gan92], an R_{TH} value for a DIBJT with $A_{tub}/A_{dev} = 22$ is extracted as 940 (°C/W). In this case, the model calculates a thermal resistance of 906 (°C/W), which gives an error of only four percent. If an error tolerance is set at 20% or less, the limit of the thermal resistance model can be evaluated; by this measure, the model can be expected to provide accurate results for DIBJT's with $A_{tub}/A_{dev} \leq 30$.

4.6 Summary

Due to the low thermal conductivity of the silicon dioxide used for the trench and buried oxide isolation, the thermal impedance of DIBJT's can be larger than that of their bulk counterparts. Therefore, self-heating effects can be enhanced and it is important to have a physical model that can predict the dynamic temperature rise. This chapter presented a thermal impedance model for DIBJT's. As with the bulk BJT/HBT's, the device interconnects can affect the thermal impedance. Neglecting the heat capacity of the emitter metallization results in a predicted thermal impedance that exaggerates the transient temperature response. Therefore, the emitter interconnect thermal impedance model, developed in Chapter Two, was utilized in this chapter. The DIBJT thermal impedance model was shown to agree reasonably well with both three-dimensional finite-element simulations and measurements. Finally, in the limit of steady-sate heat conduction, the thermal impedance model was simplified to provide a single, closed-form expression for the thermal resistance.

CHAPTER 5
A THREE-DIMENSIONAL THERMAL IMPEDANCE MODEL FOR BULK METAL-OXIDE-SEMICONDUCTOR FIELD-EFFECT TRANSISTORS

5.1 Introduction

Due to the development of complementary metal-oxide-semiconductor technologies (CMOS), the bulk MOS field-effect transistor (MOSFET) has become the primary device used in semiconductor circuits. Figure 5.1 illustrates the cross-section of a typical bulk MOSFET structure. The MOSFET output current is not as strongly dependent on temperature as the current of bipolar transistors. The current's sensitivity to changes in operating temperature is mainly due to variations in the carrier mobility. A significant change in output current will be observed only for temperature variations of tens of degrees. Based on the thermal resistances of typical devices, such temperature excursions will only occur at high power levels. However, as MOSFET's are aggressively scaled towards tenth-micron channel lengths, thermal impedances are likely to increase and self-heating effects could be enhanced.

Previous works in this area have provided methods for calculating the thermal impedance of bulk MOSFET's; however, these approaches were limited by inadequate derivations. In the work by Schutz et al. [Sch84], the two-dimensional steady-state temperature distribution in a MOSFET was solved using a finite difference discretization of the heat conduction equation. The solution was
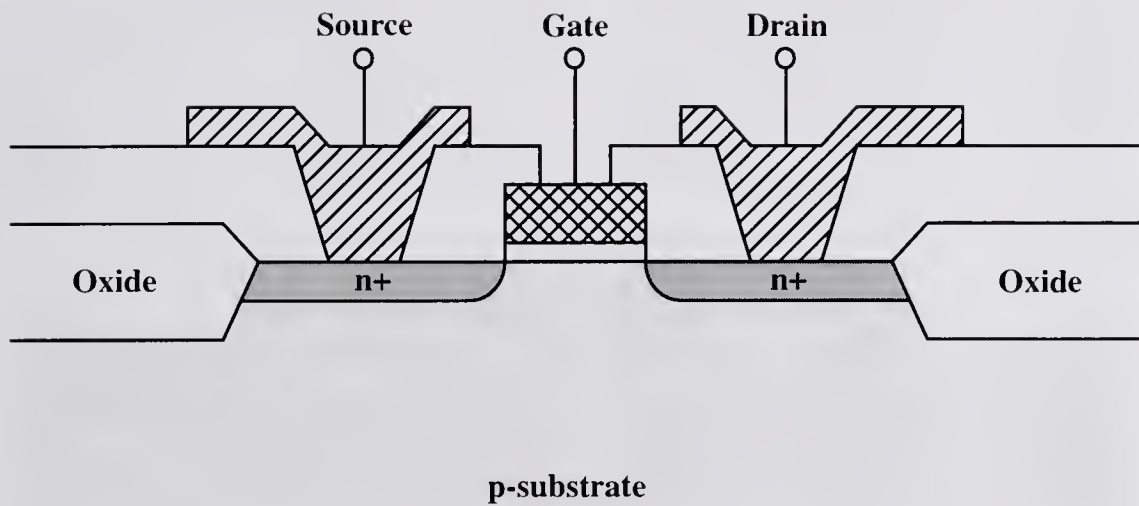
Figure 5.1   Cross-section of a typical bulk metal-oxide-semiconductor field-effect transistor (MOSFET).

calculated for a thin layer about the channel, where the boundary conditions were set by an effective one-dimensional (1-D) thermal resistance representing conduction in the substrate. The idea of the effective substrate thermal resistance was extended by Hirsch et al. [Hir93] into a quasi-three-dimensional analysis by coupling parallel 1-D analyses. While the above-mentioned techniques offer insight into steady-state heat conduction in bulk MOSFET's, they are not capable of determining the dynamic temperature response. A model for the dynamic thermal impedance of bulk MOSFET's was developed by Sharma and Ramanathan [Sha83]. However, this model was derived by neglecting the temperature variation along the width of the channel and the variation of the electric field along the length of the channel. Both of these assumptions can lead to significant errors in the predicted temperature rise within the channel region.

The model derived in this chapter is based on an extension of the model developed by Sharma and Ramanathan [Sha83], and provides a closed-form physical solution for the transient thermal impedance of bulk MOSFET's. The first part of this chapter details the derivation of the model and the improvements made to the existing work. The effects of LOCOS and shallow trench isolation and the drain, gate and source interconnects on the thermal impedance are then investigated using three-dimensional finite-element simulations. Finally, the accuracy of the thermal impedance model is evaluated using three-dimensional finite-element simulations and measurements.

5.2 Derivation of the Bulk MOSFET Thermal Impedance Model

For the derivation of the bulk MOSFET thermal impedance model, the silicon substrate is represented by a homogeneous semi-infinite half-plane with an adiabatic top surface (no heat transfer perpendicular to the surface). The back side of the substrate is assumed to be held at a constant temperature, T_0. Figure 5.2 illustrates the simplified device geometry assumed for the model derivation; the diagram focuses on the electrically active portion of the device around the drain, source and channel, which has a width W and length L. The heat source represents the power generated in the channel, which is assumed to be uniform along the channel width.

Representing the substrate as a semi-infinite medium neglects the influence of the back-side and the lateral edges of the wafer, as well as any LOCOS or trench isolation structures, on the thermal response of the device. The ramifications of neglecting the isolation structures are investigated later in the chapter. The surface of the wafer is assumed to be the only boundary that affects the thermal response of the device and it is considered to be adiabatic. With respect to the assumptions concerning the back-side and lateral edges of the substrate, the bulk MOSFET model is identical to the bulk BJT model and the validation in Chapter Two applies. Assuming that the surface of the device is adiabatic implies that conduction through the source, gate and drain interconnects and conduction through overlying oxide layers are neglected. As shown in Chapter Two, the thermal conductance from the device through the overlying oxide is approximately two to three orders of
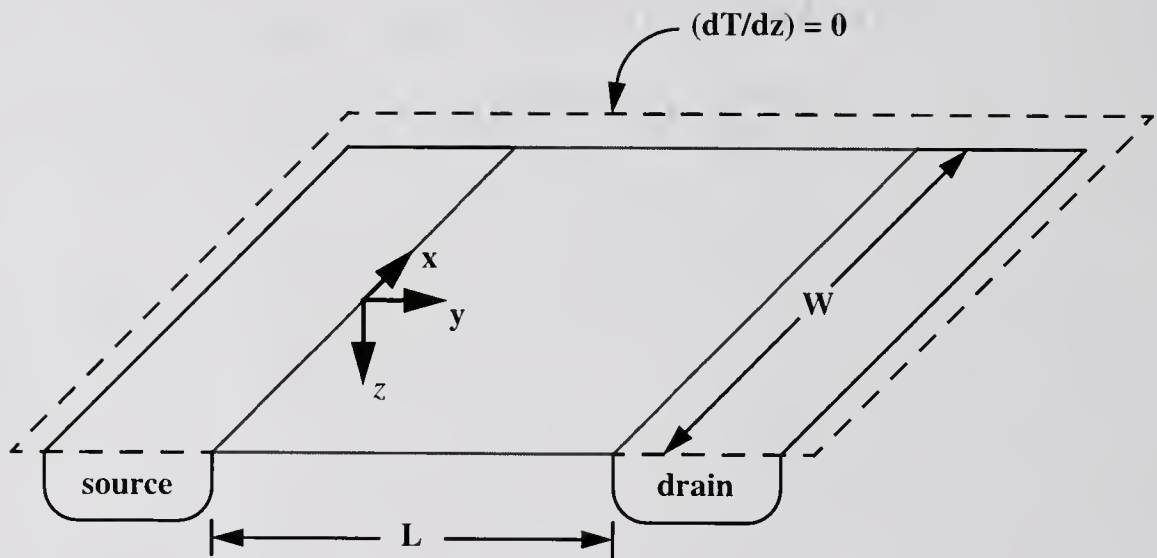
Figure 5.2   The simplified device geometry used to define the solution domain for the bulk, MOSFET thermal impedance model. The substrate is represented by a semi-infinite half-plane with an adiabatic surface (the dashed lines). Heat is generated due to the power dissipation in the channel, which has a width W and length L.

magnitude smaller than the device-to-substrate conductance. Therefore, neglecting thermal energy transport from the field regions (the regions covered with thick oxide layers) of the device is reasonable. However, the device interconnects can have high thermal conductivities, and can provide effective paths for heat flow from the portions of the device they contact. The consequences of neglecting the interconnects are investigated later in the chapter.

The heat source that represents the power dissipation in the channel depends on the region of operation and is considered to be a superposition of two individual heat sources. One source models the power dissipation in the field-dependent-velocity portion of the channel (referred to as the linear source). The other heat source models the power dissipation in the portion of the channel where the carrier velocity is saturated (referred to as the saturated source). The total temperature rise in the device can be expressed as a sum of the individual components due to the separate heat sources, and is given by

$$\Delta T(t) = Z_{THss}(t) \cdot P_{ss} + Z_{THls}(t) \cdot P_{ls}, \tag{5.1}$$

where $Z_{THss}(t)$ and $Z_{THls}(t)$ are the transient thermal impedances of the saturated and linear sources, respectively. When a MOSFET is operating in the saturation region, the power dissipation associated with the saturated source is approximately

$$P_{ss} = (V_{ds} - V_{dss})I_{ds}, \tag{5.2}$$

and for the linear source

$$P_{ls} = V_{dss}I_{ds} , \tag{5.3}$$

where V_{dss} is the drain-source voltage at the onset of the saturation region. For a device operating in the linear region, the effect of the saturated source is removed

$$P_{ss} = 0 \tag{5.4}$$

and the linear source accounts for the total power dissipation

$$P_{ls} = V_{ds}I_{ds} . \tag{5.5}$$

The original model derived by Sharma and Ramanathan assumed that the temperature gradient along the width of a MOSFET was negligible [Sha83]. However, ANSYS simulations of a MOSFET structure have shown that the temperature variations across the width of a device, for aspect ratios (W/L) greater than ten, can be in excess of 10% to 17%. Therefore, the two-dimensional analysis of the original model is extended into three dimensions. The temperature rise at any point within the device can be described by the nonhomogeneous three-dimensional heat conduction equation

$$\nabla^2 \Delta T(x, y, z, t) + \frac{g(x, y, z, t)}{k} = \frac{1}{\alpha} \frac{\partial \Delta T(x, y, z, t)}{\partial t} \tag{5.6}$$

and the following boundary conditions

$$\Delta T(\pm\infty, y, z, t) = 0 \tag{5.7}$$

$$\Delta T(x, \pm\infty, z, t) = 0 \tag{5.8}$$

$$\left.\frac{\partial \Delta T(x, y, z, t)}{\partial z}\right|_{z=0} = 0 \tag{5.9}$$

$$\Delta T(x, y, \infty, t) = 0, \tag{5.10}$$

where ΔT is the temperature rise above the local ambient ($\Delta T = T - T_0$), g is the internal energy generation density, k is the thermal conductivity, α is the thermal diffusivity ($\alpha = k/(\rho \cdot c_p)$ where ρ is the density and c_p is the specific heat) and t is time. Typical values for the material properties of bulk silicon are given in Table 5.1.

Table 5.1 Material properties of silicon

Property	Value
k (W cm^{-1} K^{-1})	1.412
ρ (g cm^{-3})	2.328
c_p (J g^{-1} K^{-1})	0.70

Source: [Mul77]

Equation (5.6) assumes that the thermal conductivity is not a function of temperature or position. Neglecting the temperature dependence of the thermal conductivity is reasonable for moderate temperature rises as compared to the ambient temperature. However, for large temperature excursions, the value of the thermal conductivity can vary significantly. The temperature dependence of k can be

accounted for using the Kirchoff transformation described in Chapter Eight. Neglecting the spatial dependence of the thermal conductivity implies that the effect of dopant atoms is ignored. In the highly doped source and drain regions, the thermal conductivity can be significantly lower than the bulk value given in Table 5.1. However, these regions are relatively small and the majority of the device substrate is relatively low-doped; therefore, as shown in Chapter Two, the average thermal conductivity of the substrate will not differ greatly from the intrinsic value.

With the initial thermal conditions in the device specified as

$$\Delta T(x, y, z, 0) = 0, \tag{5.11}$$

the solution to (5.6) can be expressed in the form

$$\Delta T(x, y, z, t) = \frac{\alpha}{k} \int_{t'=0}^{t} dt' \int_{V} G(x, y, z, t | x', y', z', t') g(x', y', z', t') dv' \tag{5.12}$$

where

$$G(x, y, z, t | x', y', z', t') = \frac{1}{8[\pi\alpha(t-t')]^{3/2}} \exp\left[\frac{-(x-x')^2}{4\alpha(t-t')}\right] \exp\left[\frac{-(y-y')^2}{4\alpha(t-t')}\right]$$

$$\cdot \left\{ \exp\left[\frac{-(z-z')^2}{4\alpha(t-t')}\right] + \exp\left[\frac{-(z+z')^2}{4\alpha(t-t')}\right] \right\} \tag{5.13}$$

is the Green's function for the given boundary-value problem [Ozi93]. Equation (5.13) physically represents the temperature at point (x, y, z) at time t, due to an instantaneous point source, g_p^i (W s), of unit strength at point (x', y', z') at time t'.

To account for a finite substrate thickness, (5.13) can be modified in a similar fashion as detailed in Chapter Two for the bulk bipolar model. The resulting expression for the Green's function is

$$G(x, y, z, t | x', y', z', t') = \frac{1}{8[\pi\alpha(t-t')]^{3/2}} \exp\left[\frac{-(x-x')^2}{4\alpha(t-t')}\right] \exp\left[\frac{-(y-y')^2}{4\alpha(t-t')}\right]$$

$$\cdot \sum_{p=1}^{\infty} \exp[-\alpha \cdot \eta_p^2(t-t')] \cdot \frac{2}{D_{sub}}$$

$$\cdot \cos(\eta_p z) \cdot \cos(\eta_p z') \tag{5.14}$$

where D_{sub} is the thickness of the substrate. The variables η_p are the eigenvalues for the boundary-value problem, and are given by the positive roots of

$$\cos(\eta_p D_{sub}) = 0. \tag{5.15}$$

As mentioned previously, the energy generation term in (5.12) is divided between two sources. The linear source is used to represent the power dissipation in the low-field portion of the channel, where the carrier velocity varies with the longitudinal electric field in proportion to the carrier mobility. The saturated source then models the power dissipation in the portion of the channel where the carrier velocity is saturated due to a large longitudinal electric field. In the linear region of operation, the low-field channel length L_e is equal to the total channel length L. For the saturated region of operation, the carrier velocity is saturated over a portion of the channel near the drain, so that $L_e < L$. In the work by Sharma and Ramanathan, the finite length of the velocity-saturated portion of the channel is neglected [Sha83]. Consequently, for all regions of operation, the linear source is modeled by a

rectangular sheet source of length L, and the saturated source is modeled by a line source at the drain/channel junction.

To produce a physically consistent model, a more rigorous analysis was attempted that models the saturated source as a rectangular sheet with a length equal to that of the velocity-saturated portion of the channel. The length of the saturated portion of the channel can be derived using the analysis by Green [Gre93], giving

$$L - L_e \cong l_c \cdot \ln\left[\frac{\mu_{eff}}{2v_{sat}l_c}\left\{(V_{ds} - V_{dss}) + \sqrt{(V_{ds} - V_{dss})^2 + 4\left(\frac{v_{sat}l_c}{\mu}\right)^2}\right\}\right], \qquad (5.16)$$

where l_c is the characteristic length of the surface potential. While (5.16) provides a more physical characterization for the length of the saturated source, two problems are associated with its implementation. First, equation (5.16) is strongly bias-dependent. Thus, the thermal impedance of the saturated source would need to be calculated as a function of device operation, rendering the thermal impedance model incompatible with the ETCS modification described in Chapter One. Second, the physical accuracy gained by using (5.16) did not translate to a significant improvement in the overall accuracy of the thermal impedance model. Therefore, the improvements offered by the use of (5.16) did not justify the complexity of its implementation, and the original approach used by Sharma and Ramanathan [Sha83] was retained.

5.2.1 The Linear Source Thermal Impedance

In the original model developed by Sharma and Ramanathan, the potential in the field-dependent-velocity portion of the channel was assumed to vary linearly along the length of the channel [Sha83]. This linear-potential assumption does not accurately represent the actual electric field in the channel. Due to the field-dependent velocity, the surface potential varies parabolically with position along the length of the channel, such that

$$\Psi_s(y) = \left(\frac{V_{ls}}{L^2}\right) y^2 , \tag{5.17}$$

where $V_{ls} = V_{ds}$ for the linear region of operation and $V_{ls} = V_{dss}$ for the saturated region of operation. Therefore, using the relation $P = I \bullet E$, the energy generation per unit-area can be expressed more accurately as

$$g_{ls}(y', t') = \left[\frac{2 \cdot P_{ls}(t')}{W \cdot L^2}\right] y' . \tag{5.18}$$

The temperature rise at any point (x, y, z) in the channel due to the linear source is calculated by substituting (5.13) and (5.18) into (5.12), and then integrating over the width and length of the channel. Assuming a step increase in power dissipation at $t' = 0$ ($P_{ls}(t') = P_{ls} \cdot U(t)$) and expressing the temperature rise as

$$\Delta T_{ls}(t) = Z_{THls}(t) \cdot P_{ls} , \tag{5.19}$$

yields the transient thermal impedance of the linear source

$$Z_{THls}(x, y, z, t) = \int_t \frac{1}{\rho c W \cdot L^2 \sqrt{\pi \alpha t}} \cdot \left[erf\left(\frac{W/2 + x}{\sqrt{4\alpha t}}\right) + erf\left(\frac{W/2 - x}{\sqrt{4\alpha t}}\right) \right]$$

$$\cdot \left\{ \sqrt{\frac{\alpha t}{\pi}} \left[\exp\left(\frac{-y^2}{4\alpha t}\right) - \exp\left(\frac{-(y - L)^2}{4\alpha t}\right) \right] \right.$$

$$\left. + y \left[erf\left(\frac{y}{\sqrt{4\alpha t}}\right) - erf\left(\frac{y - L}{\sqrt{4\alpha t}}\right) \right] \right\} \cdot \exp\left(\frac{-z^2}{4\alpha t}\right) \cdot dt \quad . \qquad (5.20)$$

Equation (5.20) represents the temperature rise at any point in the device normalized to a unit-step increase in power dissipation in the linear source. To account for a finite substrate thickness, (5.14) can be used in place of (5.13) in equation (5.12).

The effect of accounting for the parabolic variation in the surface potential is illustrated in Figure 5.3, where (5.20) is compared to a thermal impedance model derived using the uniform-field assumption of Sharma and Ramanathan [Sha83]. Since assuming a uniform electric field does not accurately model the increased power dissipation near the drain, the original model significantly under-estimates the steady-state thermal resistance in that region.

For circuit simulation, a single value is needed to represent the effective operating temperature of the device. Therefore, the thermal impedance model given by (5.20) should be evaluated at a single point. Appropriate values for coordinates can be determined by tuning the temperature rise predicted by (5.1) to measured data in a region of steady-state heat flow. Based on the extracted temperature data shown later in the chapter, the optimum point for the model evaluation was found to be $(x = 0, y = 0.95 \cdot L, z = 0)$.
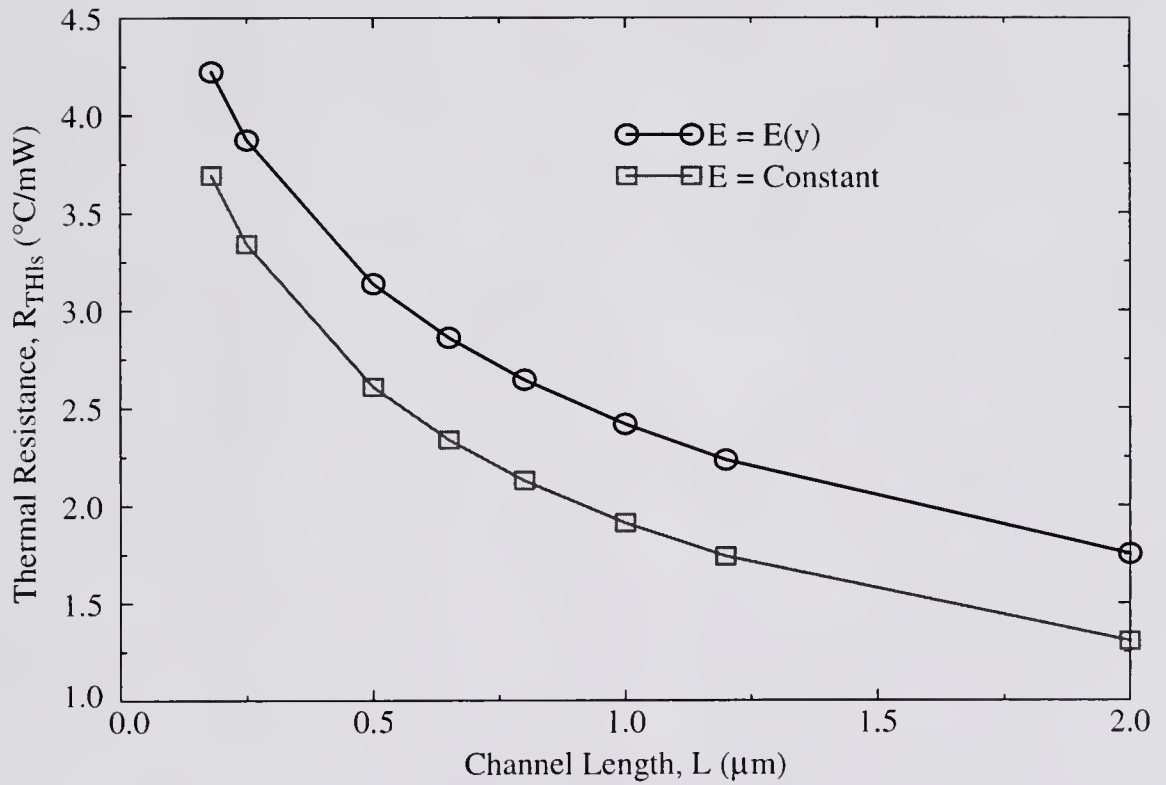
Figure 5.3   Comparison of thermal resistance calculated assuming a uniform electric field, E, to the thermal resistance calculated using (5.18), for W = 2 μm. In both cases, the thermal resistance model was evaluated at (x = 0, y = L, z = 0).

5.2.2 The Saturated Source Thermal Impedance

The heat source that represents the power dissipation in the velocity-saturated portion of the channel is modeled by a line source located at the surface of the drain/channel junction. The energy generation per unit width can therefore be expressed as

$$g_{ss}(t') = \frac{P_{ss}(t')}{W}.$$

(5.21)

The temperature rise at any point (x, y, z) in the channel due to the saturated source is calculated by substituting (5.13) and (5.21) into (5.12), and then integrating over the width of the channel. Assuming a step increase in power dissipation at $t' = 0$ ($P_{ss}(t') = P_{ss} \cdot U(t)$) and expressing the temperature rise as

$$\Delta T_{ss}(t) = Z_{THss}(t) \cdot P_{ss},$$

(5.22)

yields the transient thermal impedance of the saturated source

$$Z_{THss}(x, y, z, t) = \int_t \frac{1}{4\rho c W \pi \alpha t} \cdot \left[erf\left(\frac{W/2 + x}{\sqrt{4\alpha t}} \right) + erf\left(\frac{W/2 - x}{\sqrt{4\alpha t}} \right) \right]$$
$$\cdot \exp\left(\frac{-(y - L)^2}{4\alpha t} \right) \cdot \exp\left(\frac{-z^2}{4\alpha t} \right) \cdot dt$$

(5.23)

Equation (5.23) represents the temperature rise at any point in the device normalized to a unit-step increase in power dissipation in the saturated source. To account for a finite substrate thickness, (5.14) can be used in place of (5.13) in equation (5.12). The thermal impedance model given by (5.23) should be evaluated at a single point to

provide an effective temperature rise suitable for circuit simulation. The optimum point for the evaluation of (5.23) is the same as for the linear source and is given by $(x = 0, y = 0.95 \cdot L, z = 0)$.

5.2.3 Effects of the Device Interconnects on the Thermal Impedance

For the derivation of the bulk MOSFET thermal impedance model, the surface of the substrate is assumed to be adiabatic. In actual devices, however, the drain, source and gate are in direct contact with interconnects that are used to electrically connect different devices on a chip. For the drain and source, the interconnects are typically fabricated using aluminum metallization, where the gate is usually contacted with polysilicon. The interconnects represent additional conductance paths which can enhance the heat transfer away from a device. Since the model neglects any such heat conduction, there is a question as to whether or not the heat loss via the interconnects significantly influences the thermal response of a device.

Three-dimensional (3-D) finite-element (FE) thermal simulations of a bulk MOSFET structure, using the ANSYS software package, were performed to examine the effects of the device interconnects on the thermal impedance. Two FE models were developed to separately investigate the effects of the drain/source interconnects and the gate interconnect; these models represented simplified device structures and were constructed using similar assumptions and boundary conditions to those detailed in Chapter Two. The FE simulations tend to over predict the heat conduction through the interconnects since the device, contacts and interconnects

were considered to be in perfect thermal contact, so that any contact resistances at the material interfaces were neglected. While the FE models do not truly represent and actual device structure, they serve as an order-of-magnitude estimate for the effects of the interconnects on the thermal response.

Figure 5.4a shows the FE model used for investigating the effects of the drain/source interconnect metallization on the thermal resistance. Steady-state thermal simulations were run with the interconnects present and with the interconnects removed. The results of the two groups of simulations were compared to determine the extent of the effect of the drain/source interconnects. The simulations were performed at various channel-to-contact spacings (L_d), for a fixed interconnect thickness (d_{met}) and a fixed oxide thickness (d_{ox}) between the interconnects and the substrate; the results are shown in Figure 5.4b. Overall, the effect of the drain/source interconnect metallization is to reduce the thermal resistance of a device; though, for the simulated devices, this effect is relatively minor. However, the data does illustrates how the drain/source interconnects become more effective thermal conductance paths as they are moved closer to the channel, or as their cross-sectional area is increased. Therefore, the effects of the drain/source interconnects on the thermal resistance of highly-scaled MOSFET's could be more significant.

The effects of the gate interconnect on the thermal resistance of bulk MOSFET's are examined by performing steady-state thermal simulations with the FE model shown in Figure 5.5a, both with and without the interconnect present. The gate material was assumed to be heavily-doped polysilicon such that the width of the
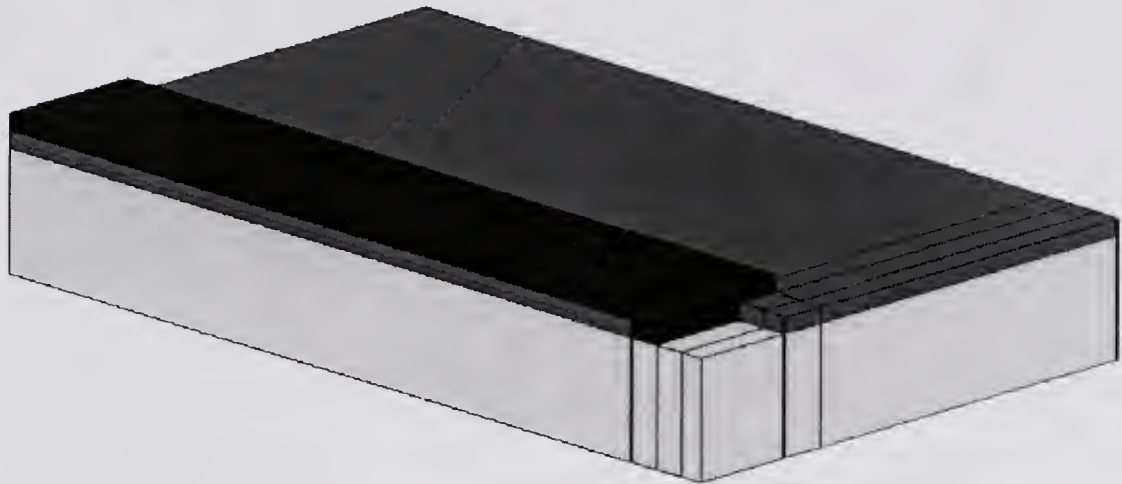
(a)
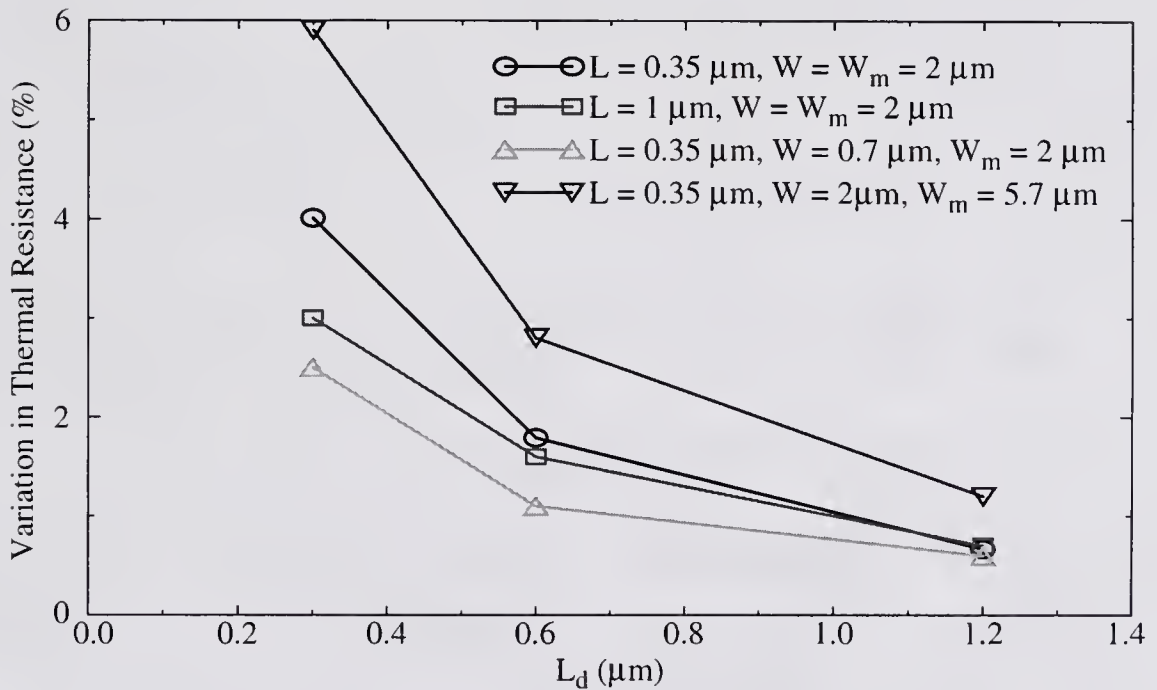
(b)

Figure 5.4   ANSYS simulations showing the effect of the drain/source interconnects on the thermal resistance, for $d_{met} = 0.7\ \mu m$ and $d_{ox} = 0.7\ \mu m$: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for the drain/source interconnects and the thermal resistance neglecting the drain/source interconnects, plotted as a function of the edge-to-edge distance between the channel and the contact openings.
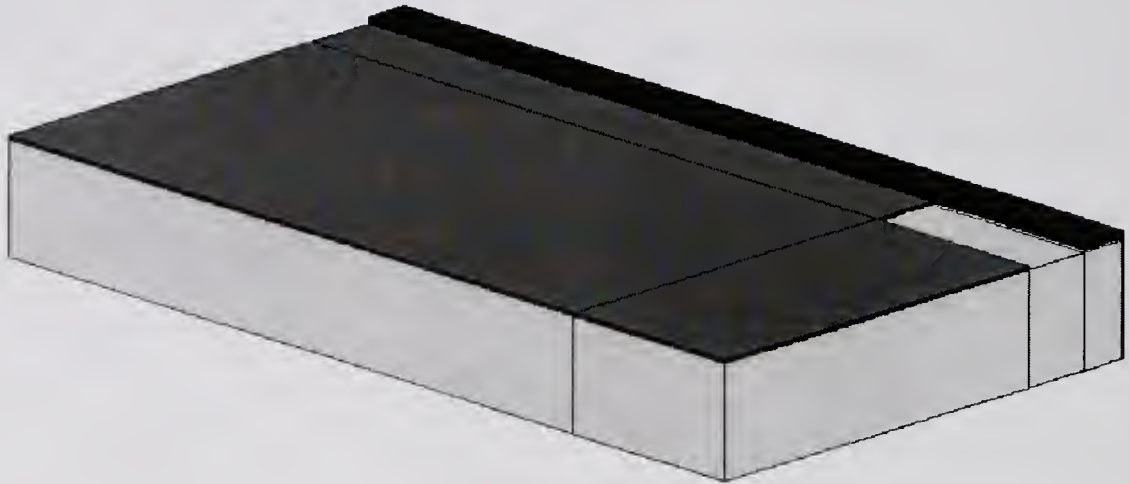
(a)
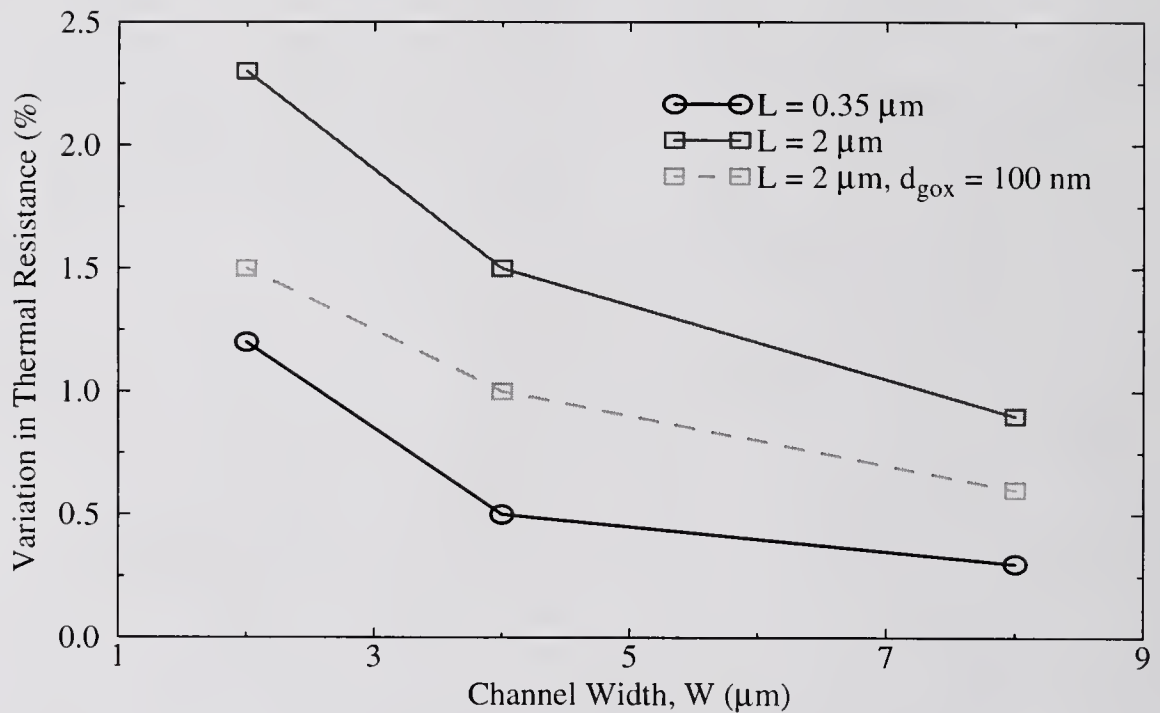
(b)

Figure 5.5   ANSYS simulations showing the effect of the gate interconnect, assumed to be heavily-doped polysilicon, on the thermal resistance, for $d_g = 0.3$ μm, $d_{gox} = 10$ nm and $d_{ox} = 0.15$ μm: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for the gate interconnect and the thermal resistance neglecting the gate interconnect, plotted as a function of the channel width.

interconnect was equal to the length of the channel. The simulations were run at various channel widths, for both short- and long-channel devices with fixed interconnect (d_g) and gate-oxide (d_{gox}) thicknesses; the results are shown in Figure 5.5b. As with the drain/source interconnects, the effect of the gate interconnect is a reduction in the device thermal resistance. The trends in the data show that this effect decreases for devices with large channel widths; and, at a fixed channel width, the reduction in the thermal resistance is more significant for longer channel lengths. However, due to the relatively low thermal conductivity of the gate oxide and heavily-doped polysilicon, the extent of the effect is only moderate. Consequently, the use of silicided gate materials and ultra-thin gate oxides, which effectively enhance the thermal conductance of the gate interconnect, could result in a more substantial reduction in the overall device thermal resistance.

Transient thermal simulations of the FE models in Figure 5.4a and Figure 5.5a were performed to examine the effect of the device interconnects on the dynamic temperature response. Figure 5.6 shows the results of the simulations for a step-increase in power dissipation. For the simulation with the drain/source interconnects, the thermal response does not deviate from that of a device without interconnects until the heat generated in the channel reaches the contacts. At this time, the additional heat capacity of the interconnect metallization effectively reduces the thermal response. The time it takes for the heat to reach the drain/source contacts can be approximated as the square of the distance L_d divided by the thermal diffusivity of the substrate material. For the simulated device structure, the resulting time is approximately one nanosecond, which agrees with the FE simulations. The

(a)

(b)
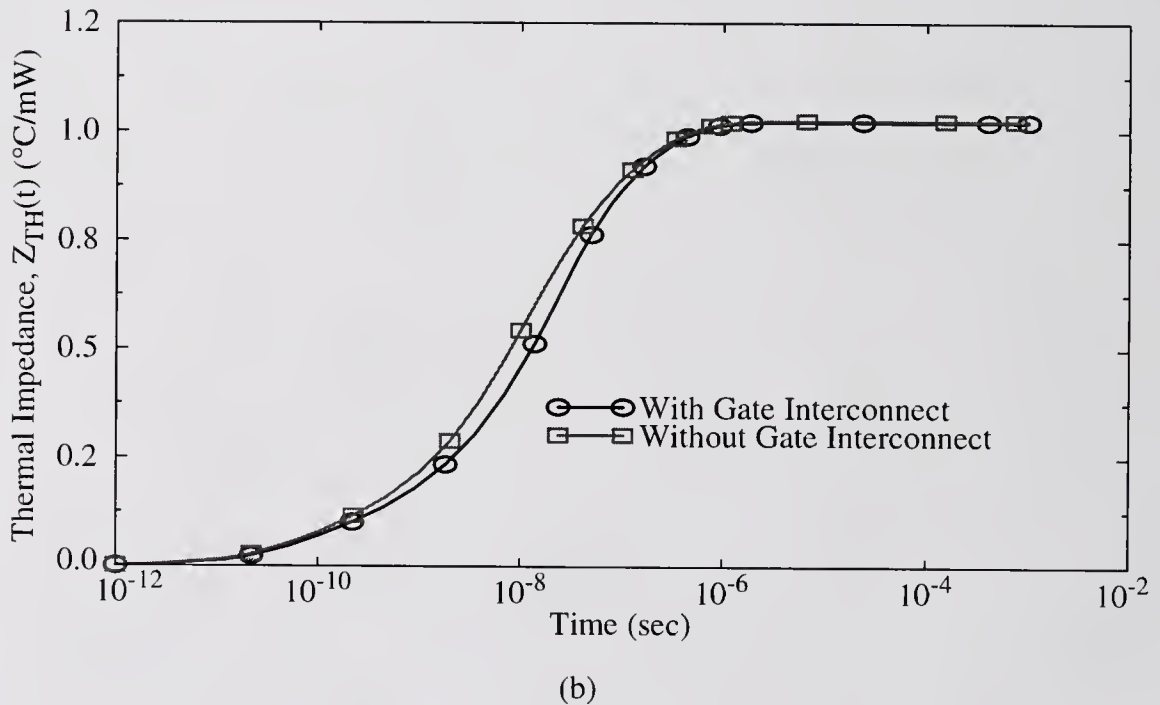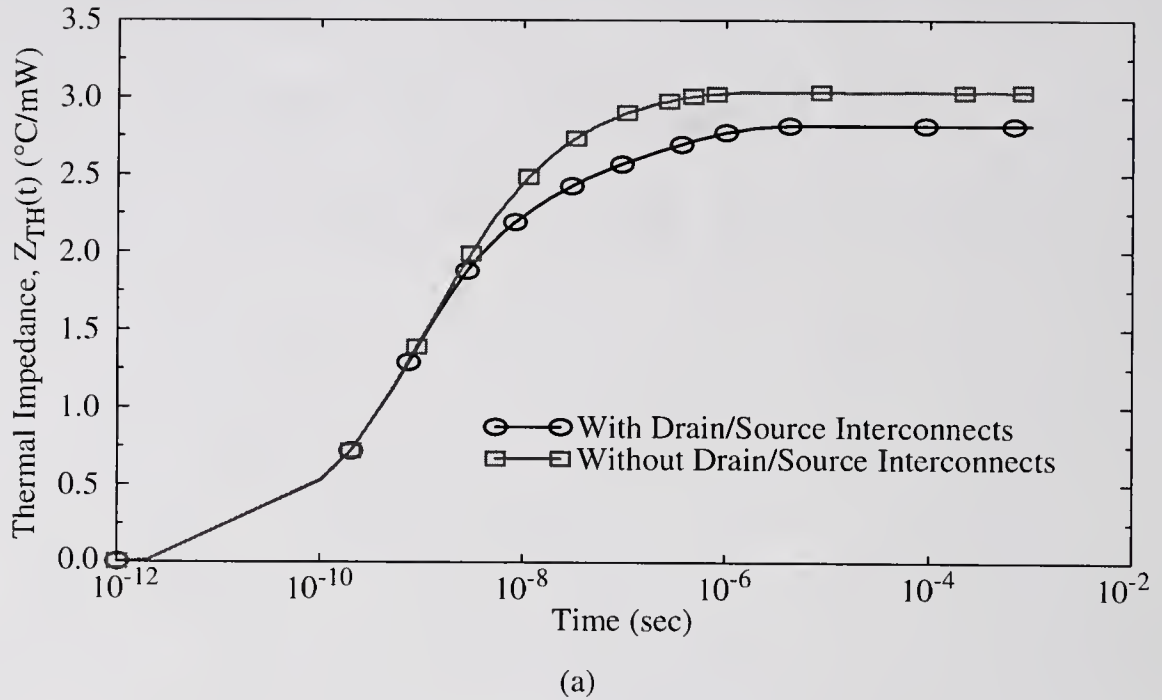
Figure 5.6   ANSYS simulations showing the effects of the device interconnects on the transient thermal impedance: a) For the drain and source interconnects with $W = 2 \ \mu m$, $L = 0.35 \ \mu m$, $d_{ox} = 0.7 \ \mu m$, $d_{met} = 0.7 \ \mu m$, $W_m = 5.7 \ \mu m$ and $L_d = 1.2 \ \mu m$; b) for the gate interconnect with $W = 4 \ \mu m$, $L = 2 \ \mu m$, $d_{ox} = 0.15 \ \mu m$, $d_{gox} = 10 \ nm$ and $d_g = 0.3 \ \mu m$.

gate interconnect affects the thermal response almost instantaneously since it is located directly above the channel. As with the drain/source interconnects, the gate interconnect represents an additional conductance path for heat to travel away from the channel, and its heat capacity reduces the thermal response of a device. However, since the thermal conductivity of heavily-doped polysilicon is relatively low, the effect of the gate interconnect on the dynamic temperature response is not substantial.

Based on the results of the 3-D FE simulations, neglecting the effects of the device interconnects in the model derivation should not result in substantial errors for larger MOSFET's. However, the effects of the interconnects on the thermal impedance can become more severe for highly-scaled devices. In such a case, the thermal impedance model will tend to over-predict both the steady-state thermal resistance and the transient rise of the thermal impedance.

5.2.4 Effects of Isolation Structures on the Thermal Impedance

As shown in Figure 5.1, a bulk MOSFET is typically surrounded by a region of silicon-dioxide which is used to electrically isolate it from neighboring devices. Such isolation is especially important for reducing latchup in CMOS circuits. The isolation technologies commonly used in bulk MOS fabrication are semi-recessed LOCOS (local oxidation of silicon) and shallow trench (e.g. BOX isolation) [Wol90]. Since the isolation structures use SiO_2, which has a low thermal conductivity, they will naturally impede the flow of heat away from a device. The bulk MOSFET thermal impedance represents the substrate as a homogeneous

material, and therefore neglects the effects associated with oxide isolation. To understand the limitations associated with this assumption, the effects of isolation structures on the thermal response should be investigated.

Three-dimensional (3-D) finite-element (FE) thermal simulations, using ANSYS, were performed to examine the effects of isolation structures on the thermal impedance of bulk MOS transistors. A single FE model was developed to investigate the effects of both LOCOS and shallow-trench isolation. The full thickness (d_{fox}) of the isolation oxide was assumed to be recessed beneath the surface of the substrate. The model represented a simplified device structure and was constructed using similar assumptions and boundary conditions to those detailed in Chapter Two. The side-walls of the isolation structures were assumed to be perpendicular to the top surface of the substrate, and were directly adjacent to the active device region. The implications of assuming vertical isolation walls are detailed in Chapter Two, where it is shown that the isolation structure model should over estimate the effects of the isolation. Therefore, while the finite-element model does not truly represent the physical device layout, it provides an order-of-magnitude estimate for the effects of the isolation structures on the thermal impedance.

Figure 5.7a shows the FE model used for investigating the effects of the isolation structures on the thermal resistance. Steady-state thermal simulations were run for various device-isolation spacings (L_{dox}); this spacing corresponds to the distance between the channel and the edge of the isolation structure, along the length of the device. Simulations were also run for the same devices with the isolation structures removed; Figure 5.7b compares the results of the FE simulations. The
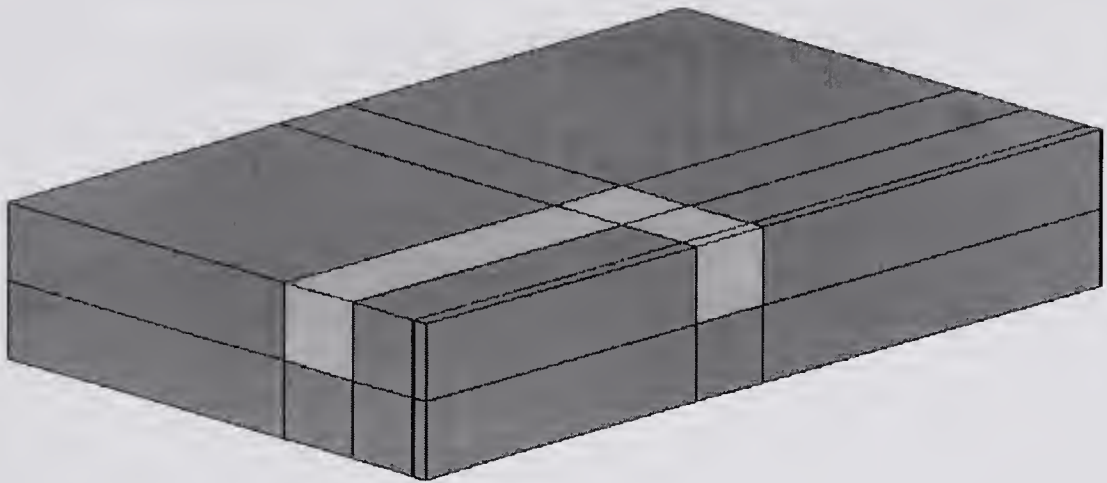
(a)
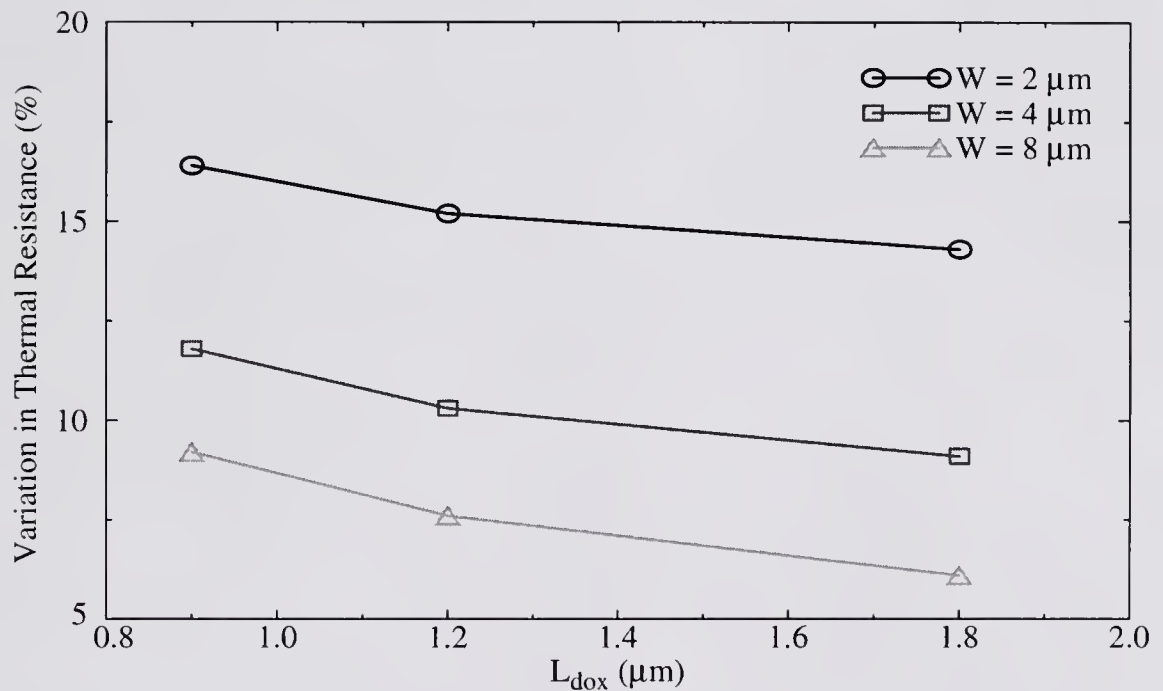
(b)

Figure 5.7   ANSYS simulations showing the effect of LOCOS or shallow trench isolation on the thermal resistance. The device specifications are $L = 0.35\ \mu m$ and $d_{fox} = 0.4\ \mu m$: a) The finite-element model simulated with ANSYS; b) the variation between the thermal resistance accounting for the isolation and the thermal resistance assuming a homogeneous substrate, plotted as a function of the distance between the channel and the isolation (along the length of the device).

oxide used in the isolation structures impedes the lateral flow of heat away from the device, resulting in an increased temperature rise over the device simulated with no isolation. The simulation data shows that this effect can be significant and increases for small-geometry devices with small device-isolation spacings. The increase in temperature rise can be reduced by decreasing the depth of the isolation structures or by moving the isolation away from the channel.

Transient thermal simulations of the FE model shown in Figure 5.7a were performed to examine the effects of both LOCOS and shallow-trench isolation on the transient thermal impedance. ANSYS was used to simulate the device with and without the isolation structure present; the results are shown in Figure 5.8. The transient temperature response is affected in the same manner as the steady-state response, and as the heat travels laterally and reaches the edges of the isolation structure, the temperature rise accounting for the isolation begins to increase over the response without the isolation.

Based on the results of the 3-D FE simulations, neglecting the effects of the isolation structures used in bulk MOSFET technologies can be considered reasonable for larger devices. However, for highly scaled devices where the effective device-isolation separation is small, the model will tend to considerably under-predict both the steady-state thermal resistance and the transient rise of the thermal impedance.
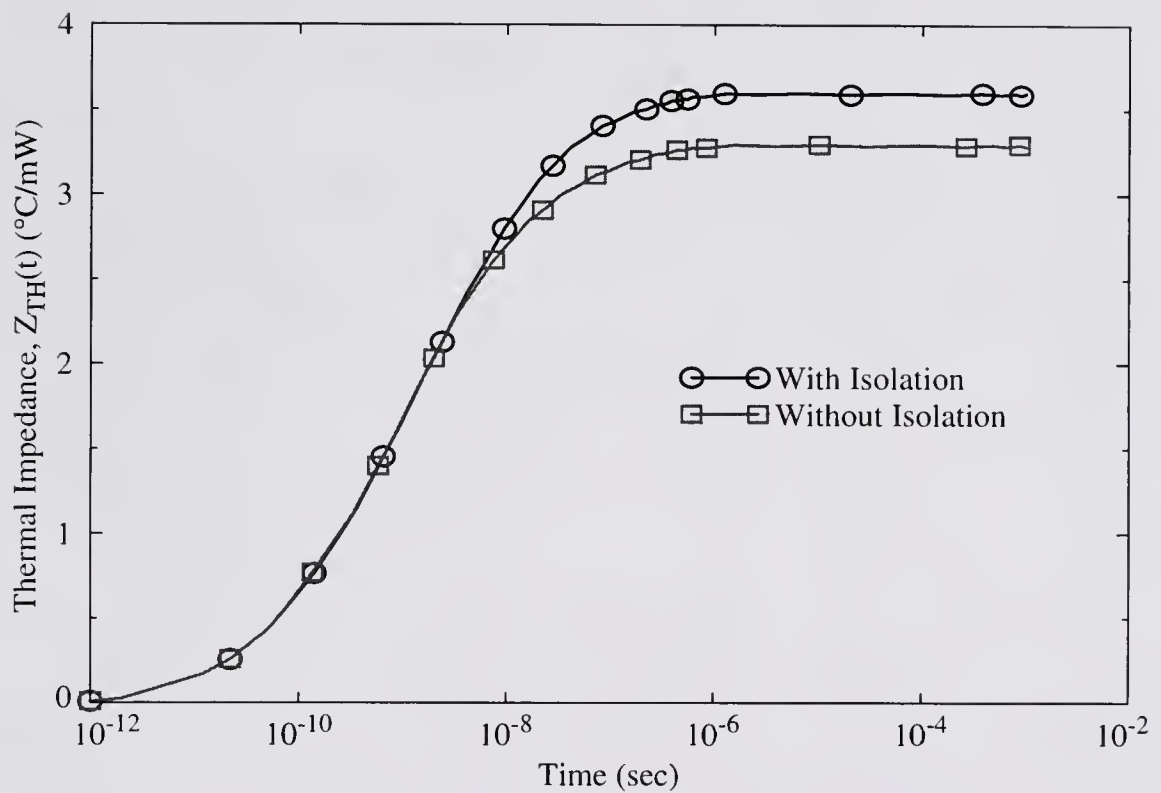
Figure 5.8   ANSYS simulations showing the effect of LOCOS or shallow trench isolation on the transient thermal impedance. The specifications for the device structure are   W = 2 μm,   L = 0.35 μm, d_{fox} = 0.4 μm, L_{dox} = 1.2 μm.

5.3 Verification of the Bulk MOSFET Thermal Impedance Model

To verify the ability of the thermal impedance model to predict the steady-state thermal resistance of bulk MOSFET's, the model was compared to measured temperature data. The temperature rise in the measured devices was predicted by evaluating the model equations (5.20) and (5.23) as $t \to \infty$ and then applying them to (5.1). The measured data were extracted using a gate-resistance thermometry technique [Goo95]. Figure 5.9, Figure 5.10 and Figure 5.11 show the comparisons between the predicted and measured values for the temperature rise. The thermal impedance model accurately predicts the steady-state temperature rise, and hence the thermal resistance, for both short- and relatively long-channel MOSFET's; the average error between the model and the measurements was 11% and 3% for the one-volt and two-volt gate biases, respectively.

The thermal impedance model was also compared to three-dimensional (3-D) transient finite-element (FE) simulations performed using ANSYS. verification. Two FE models were developed to separately verify the linear and saturated source transient impedances. To simplify the FE simulations, the MOSFET structures were assumed to be symmetric about the y-axis (see Figure 5.2); therefore, only one-half of a device was simulated. The device interconnects and device isolation were neglected and the substrate was assumed to be silicon with the bulk properties given in Table 5.1. The FE simulations were compared to the transient thermal impedance models given by (5.20) and (5.23), using the modification for finite substrate thickness; the results are shown in Figure 5.12 and Figure 5.13. The models for both the linear and saturated sources accurately replicate the transient and steady-state
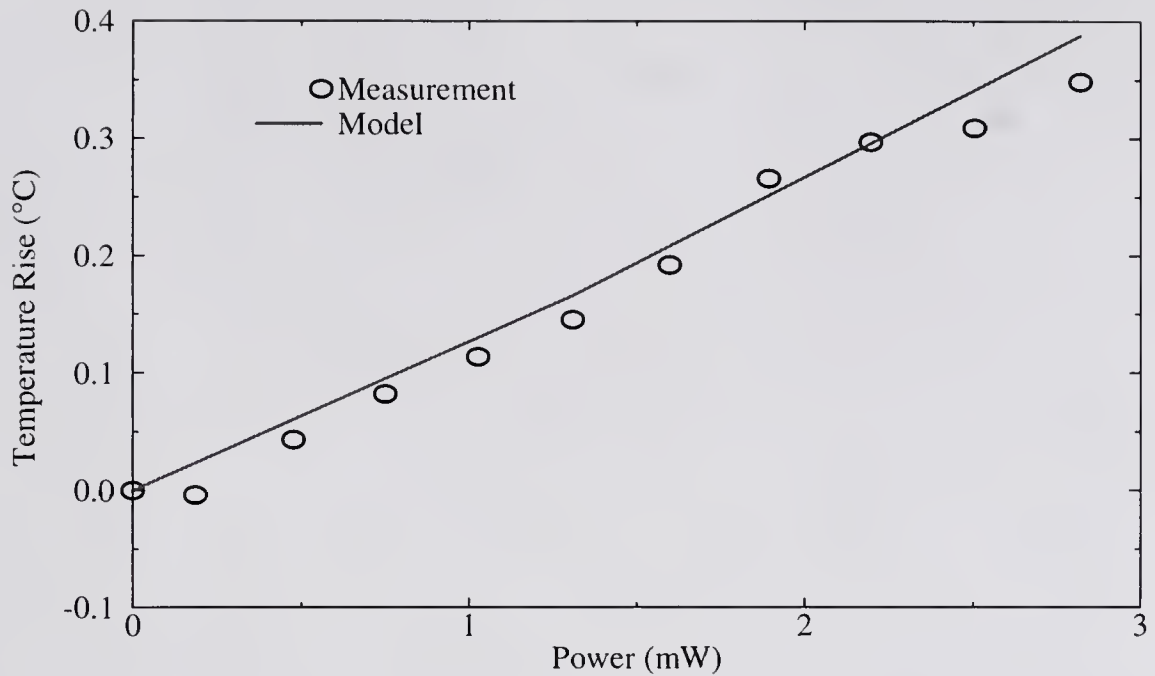
(a)
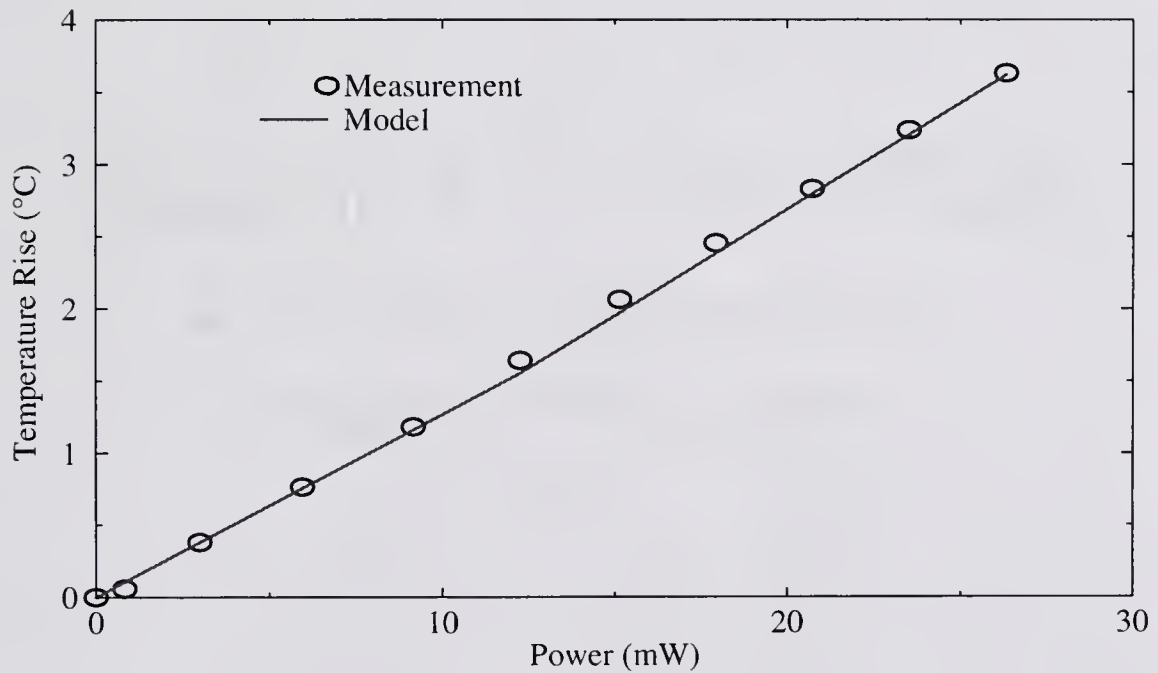
(b)

Figure 5.9   Temperature rise, extracted using gate resistance thermometry, in Motorola bulk MOSFET's with W = 120 μm and L = 0.65 μm. The drain voltage was swept from 0.0 to 2.0 V: a) 1.0 V on the gate; b) 2.0 V on the gate.
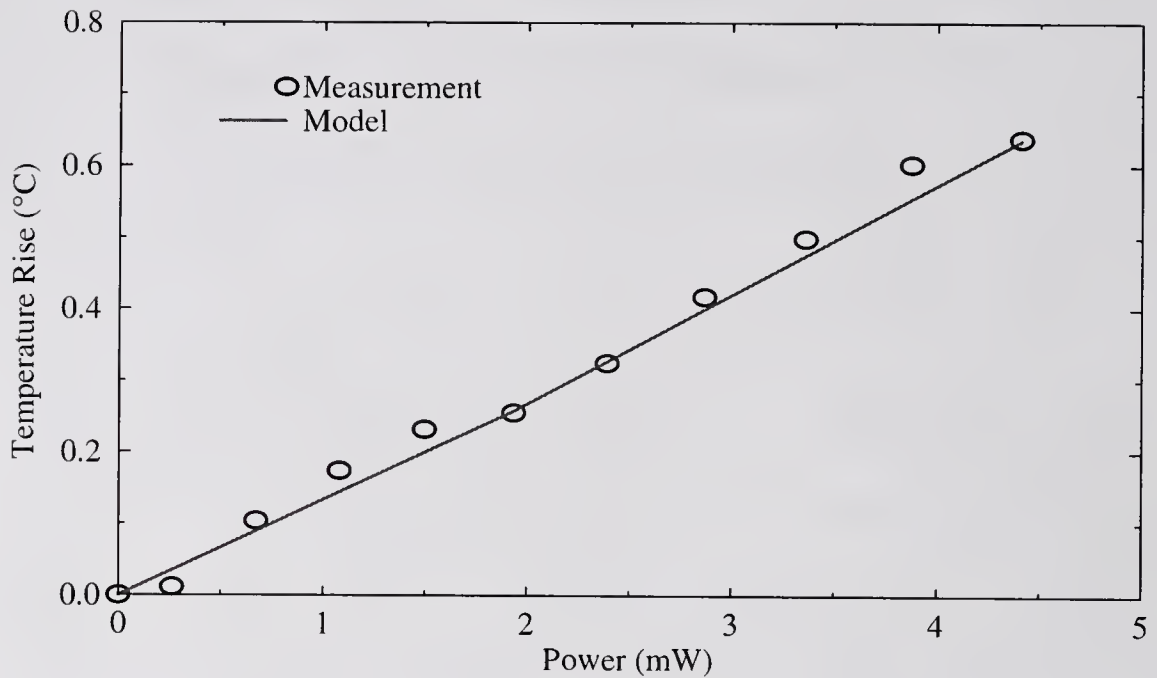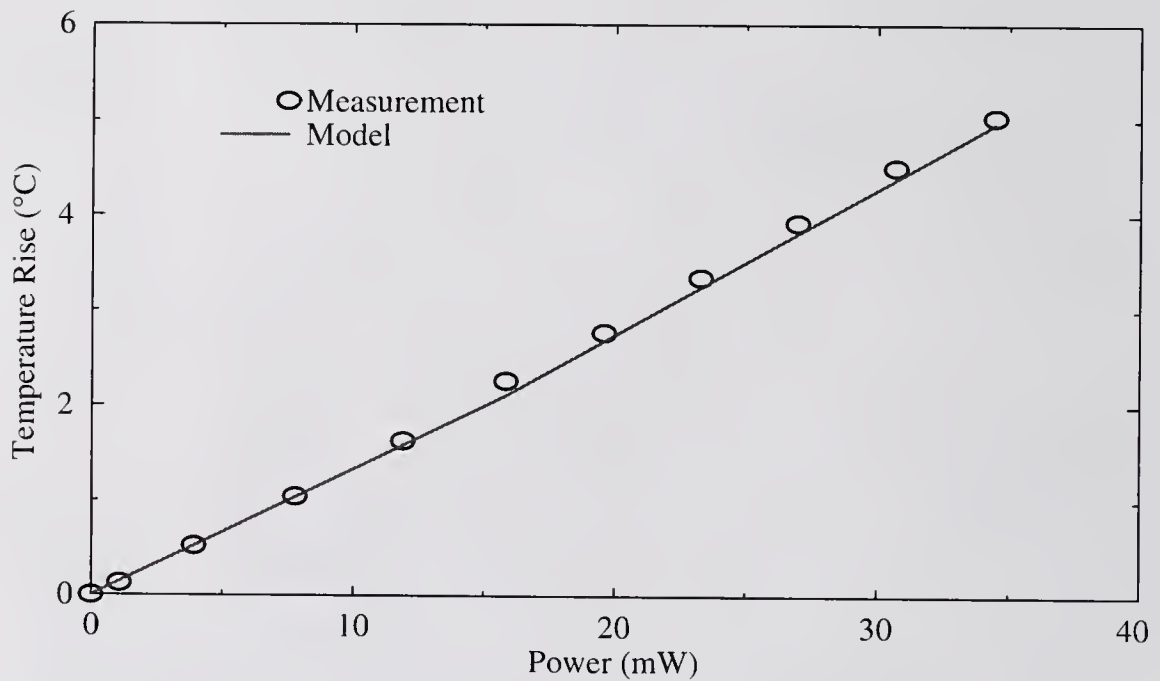
Figure 5.10   Temperature rise, extracted using gate resistance thermometry, in Motorola bulk MOSFET's with W = 120 μm and L = 0.45 μm. The drain voltage was swept from 0.0 to 2.0 V: a) 1.0 V on the gate; b) 2.0 V on the gate.
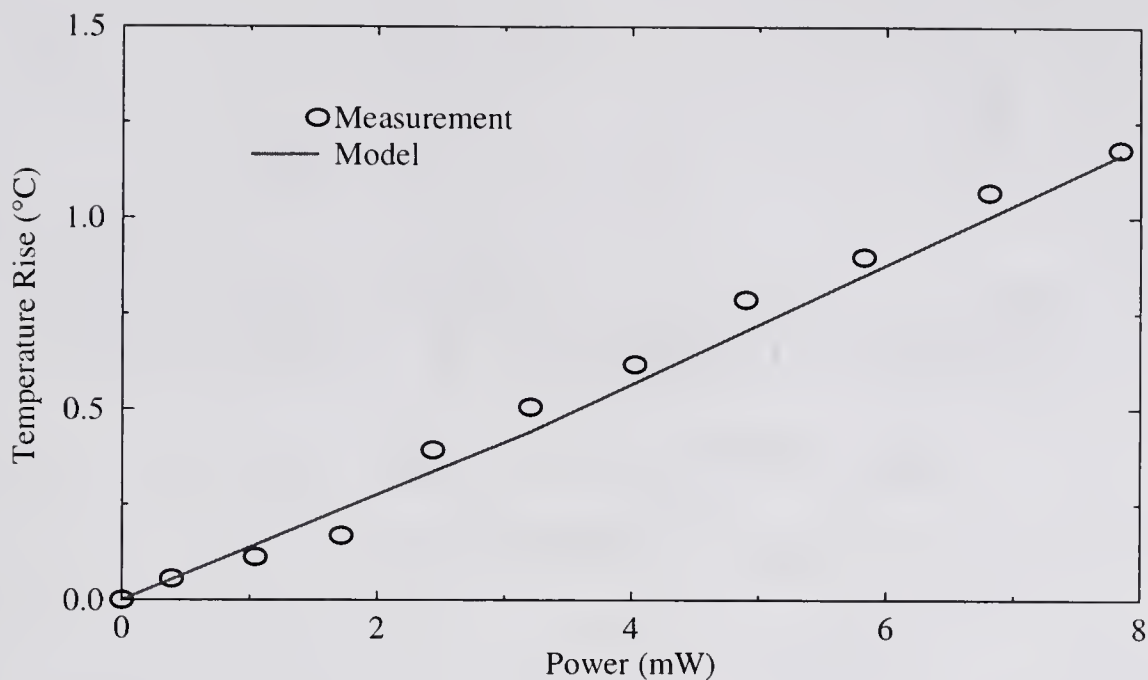
(a)
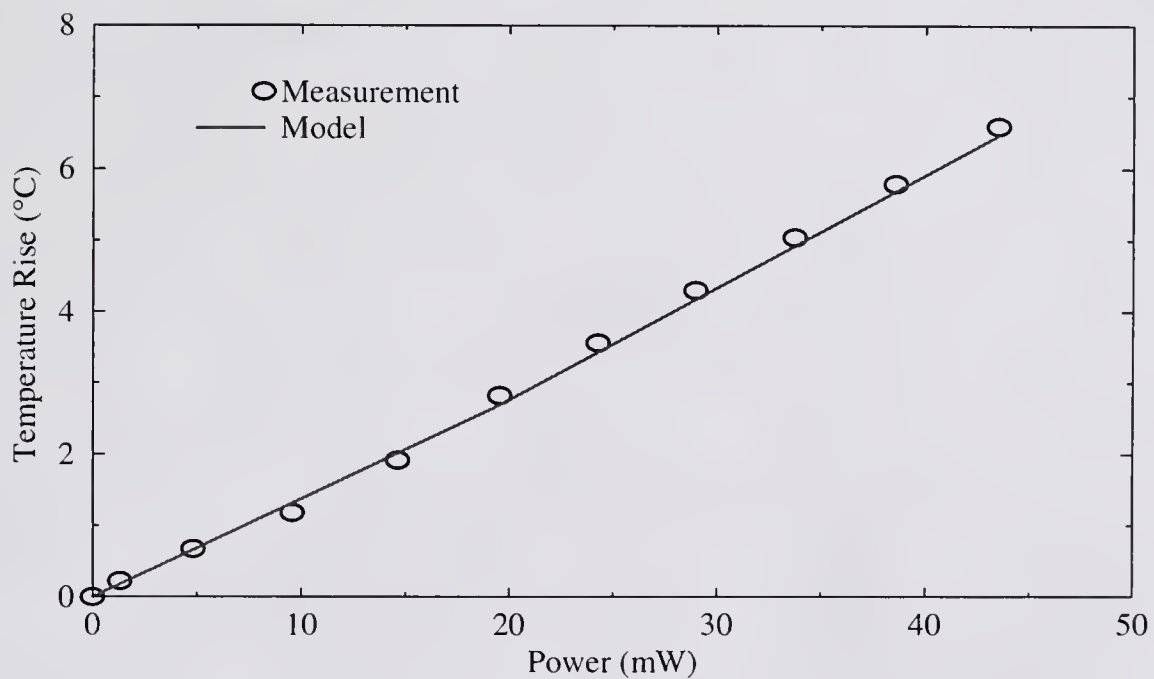
(b)

Figure 5.11   Temperature rise, extracted using gate resistance thermometry, in Motorola bulk MOSFET's with W = 120 µm and L = 0.35 µm. The drain voltage was swept from 0.0 to 2.0 V: a) 1.0 V on the gate; b) 2.0 V on the gate.
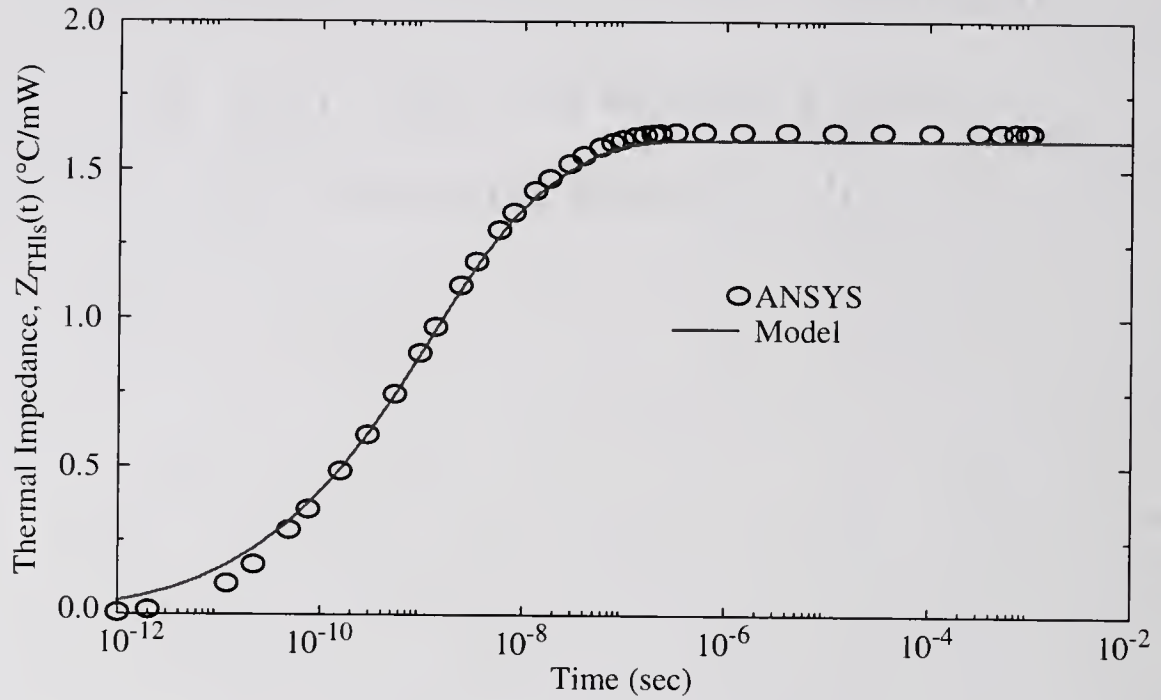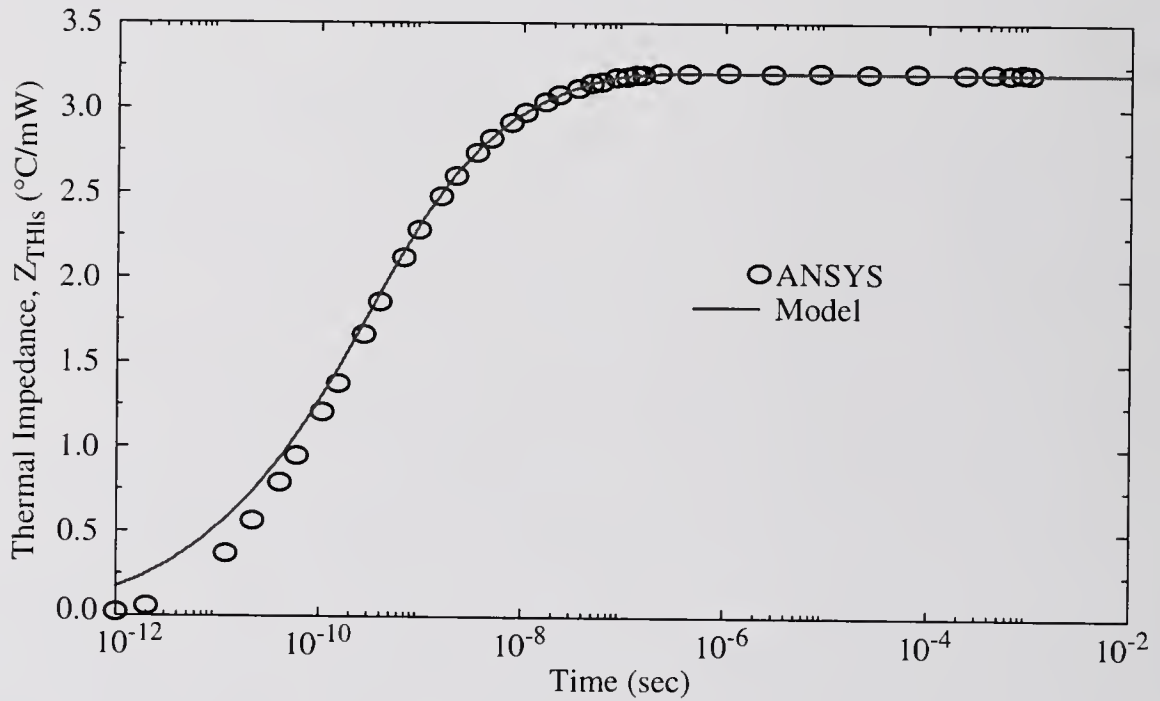
Figure 5.12   The transient thermal impedance of the linear source simulated with ANSYS and calculated with the model: a) For W = 2 μm, L = 0.65 μm; b) for W = 1 μm, L = 0.35 μm.
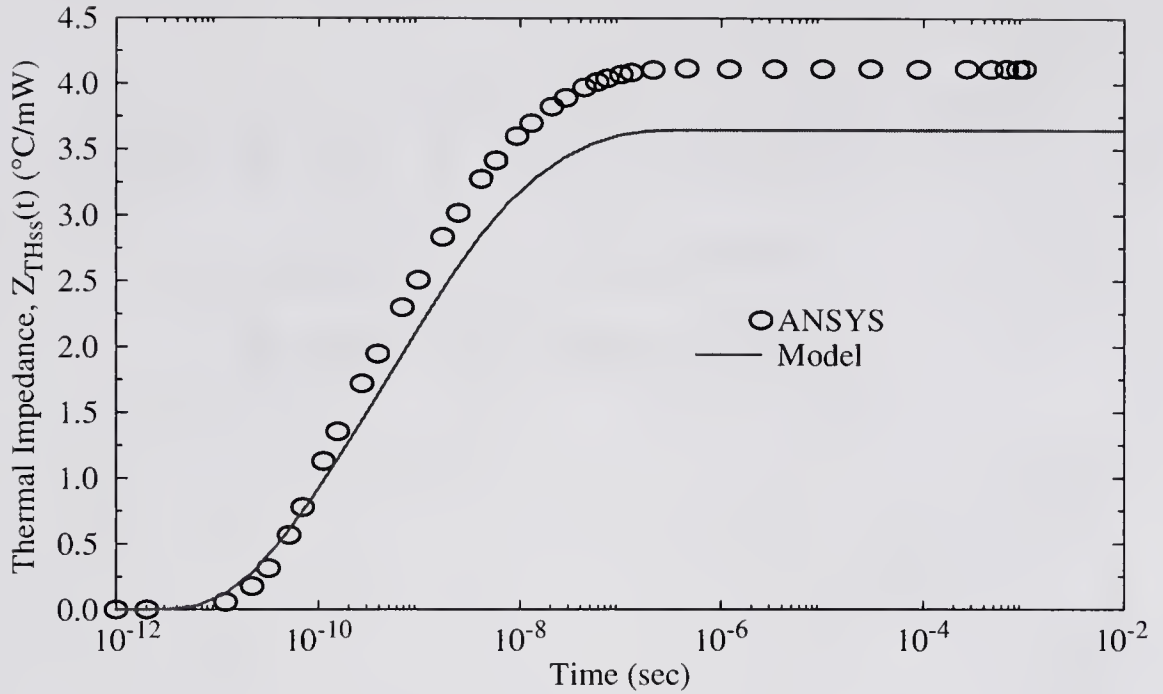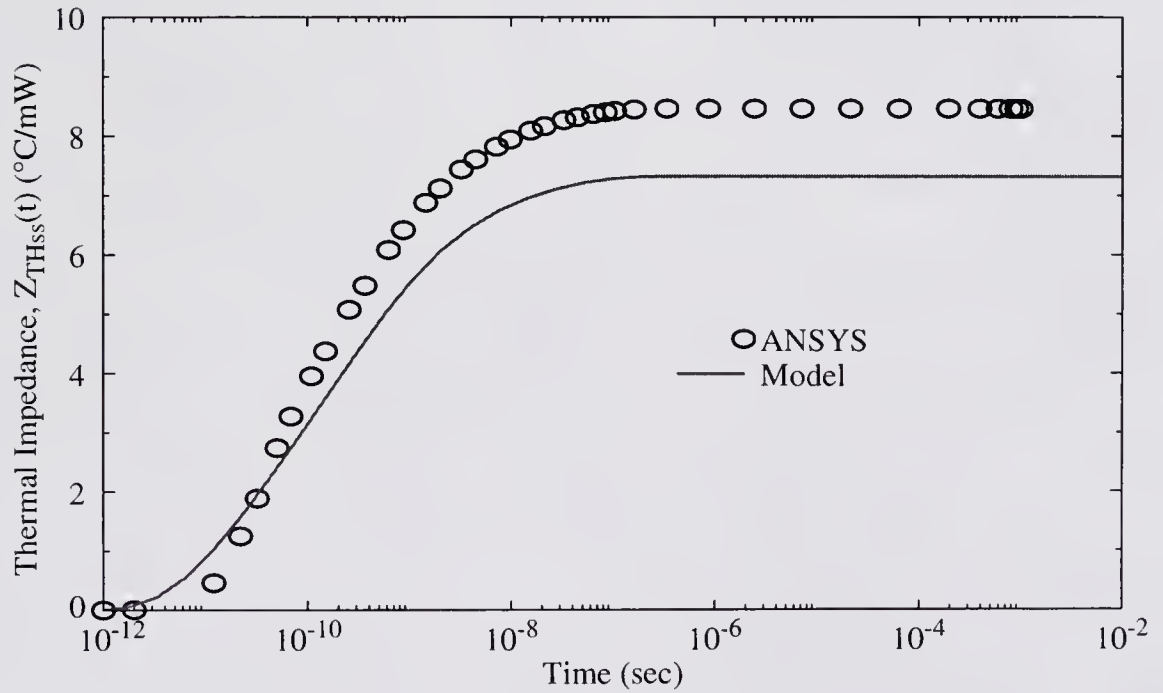
Figure 5.13   The transient thermal impedance of the saturated source simulated with ANSYS and calculated with the model: a) For W = 2 μm, L = 0.65 μm; b) for W = 1 μm, L = 0.35 μm.

temperature response; for the device structures that were simulated, the error between the FE and analytical models is no greater than 13%.

Since the parameters for the bulk MOSFET thermal impedance model are based solely on the device geometry, the model can be applied to predict the expected trends for self-heating in highly-scaled devices. Figure 5.14 shows the thermal resistance, of both the linear and saturated sources, plotted as a function of channel length. The model predicts that as bulk MOSFET's are scaled beyond quarter-micron channel lengths, a substantial increase in the thermal resistance can be expected. Therefore, self-heating could become a significant problem in future generation devices and circuits.

5.4 Summary

A thermal impedance model for bulk MOSFET's was derived. The model was shown to agree reasonably well with three-dimensional finite-element simulations and measurements of bulk devices. Since the model does not account for interconnect metallization and either LOCOS or trench isolation, the effects of neglecting such structures were investigated. The results suggest that the model can be expected to provide reasonable predictions for the thermal impedance of larger bulk devices. However, for highly-scaled devices, the effects of the drain, gate and source interconnects and isolation structures can be significant and the accuracy of the model will decline. Methods for modeling the effects of isolation structures are proposed in Chapter Eight.
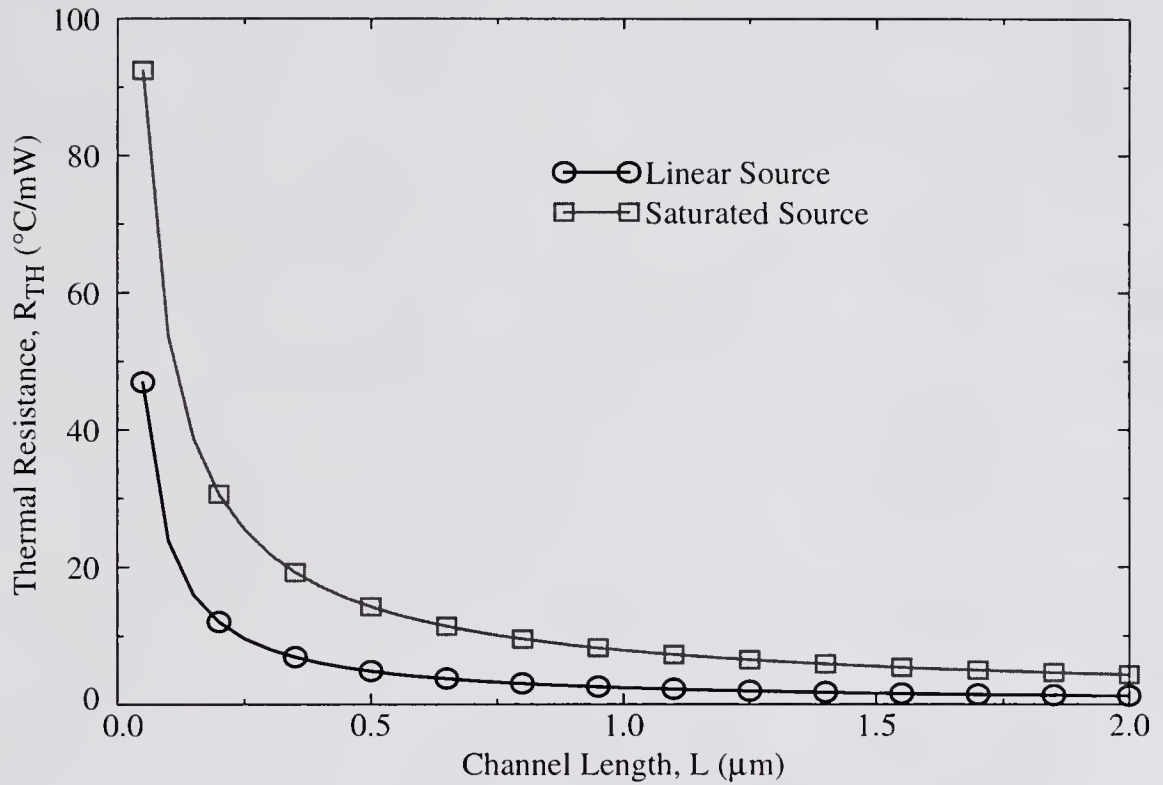
Figure 5.14   Thermal resistance calculated using the model evaluated at steady-state, plotted as a function of channel length for W/L = 2.

CHAPTER 6
A QUASI-THREE-DIMENSIONAL THERMAL IMPEDANCE MODEL FOR SILICON-ON-INSULATOR METAL-OXIDE-SEMICONDUCTOR FIELD-EFFECT TRANSISTORS

6.1 Introduction

Metal-oxide-semiconductor field-effect-transistors (MOSFET's) fabricated using silicon-on-insulator (SOI) substrates offer many advantages over MOSFET's fabricated using bulk wafer technologies. Device speed can be enhanced due to the reduction of the parasitic capacitances to the substrate [Tu94, Goo95]. The improved device isolation prevents latchup in complementary MOSFET (CMOS) circuits since the buried oxide layer eliminates the parasitic bipolar devices between transistors [Wol90]. Short-channel effects can also be reduced due to the limited vertical depletion depth imposed by the finite silicon film thickness [Wol90]. Finally, the use of SOI substrates improves radiation hardness and can increase the device packing density [Wol90].

SOI MOSFET's are commonly fabricated using SIMOX (separation by implanted oxygen) SOI wafers. Figure 6.1 shows a diagram of a typical SOI MOSFET. SIMOX wafers are created by implanting oxygen ions into a silicon substrate. The implant is followed by a high-temperature anneal in an inert ambient to form silicon dioxide (SiO_2) [Wol90]. The oxygen is implanted far enough into the silicon such that the SiO_2 is buried beneath the wafer surface; typical values for the
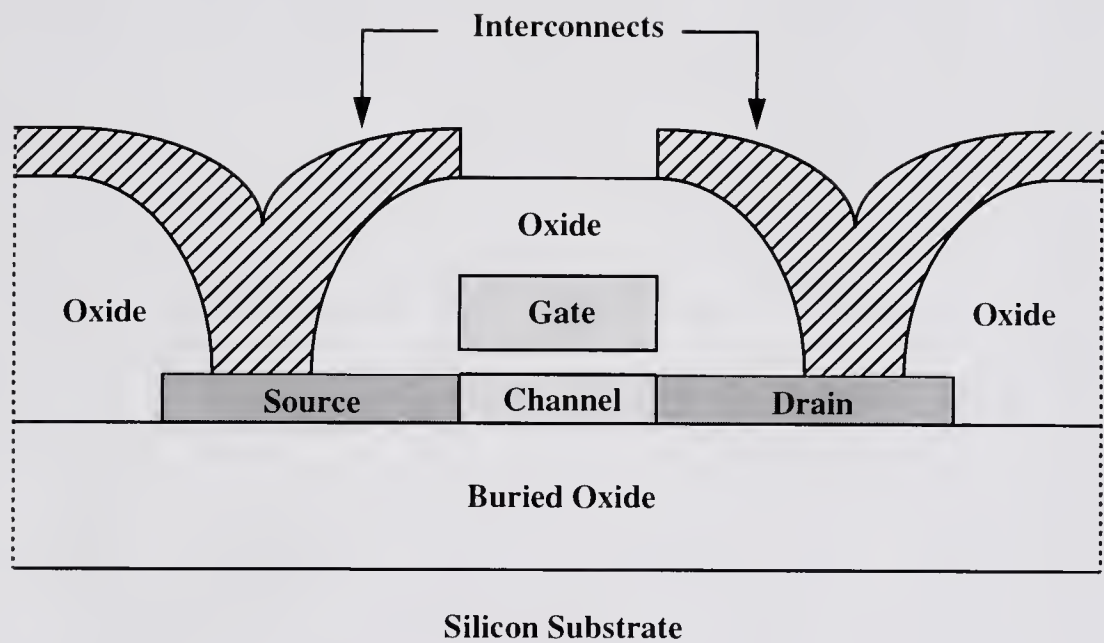
Figure 6.1   Cross-section of a typical silicon-on-insulator (SOI) metal-oxide-semiconductor field-effect transistor (MOSFET)

thickness of the buried oxide layer range from 0.2 to 0.5 μm [Tu94, Goo95, Lee95]. Once the buried oxide layer is formed, the surface silicon layer is used for the fabrication of the MOSFET's, which can follow the standard process steps used for bulk devices. Typical values for the thickness of the silicon film range from 0.04 to 0.2 μm [Tu94, Goo95, Lee95].

Due to the low thermal conductivity of the buried oxide layer, MOSFET's fabricated on SOI wafers exhibit enhanced self-heating effects [Ber91, Che95, Jom95a]. Empirical extraction techniques using gate resistance thermometry or output conductance measurements have been developed to characterize the steady-state thermal resistance of SOI MOSFET's [Goo95, Lee95, Ten95]. The dynamic thermal impedance has also been extracted using output admittance or transient drain current measurements [Cav93, Cav95, Lee95]. The shortcomings of these methods are that they require complicated measurements or the fabrication of special test structures, and they do not provide predictive values for the thermal resistance or impedance. Physics-based models have generally been limited to the steady-state thermal resistance [Ber91, Goo92, Che95, Jom95b]. A physical dynamic electrothermal model for SOI MOSFET's was developed by Bielefeld et al. [Bie95]. However, this model is solved using complex numerical techniques making it inefficient for circuit simulation.

This chapter details the derivation of a compact physics-based model for the dynamic thermal impedance of SOI MOSFET's. The thermal impedance model is an extension of a modified version of the steady-state thermal resistance model presented by Goodson and Flik [Goo92]. The work by Goodson and Flik provides an

accurate model for the thermal resistance of SOI MOSFET's for which the width of the drain/source interconnect metallization approximately equals the width of the device, and the device interconnects play a significant role in heat evacuation. However, for many analog and some digital applications, the width of the device is greater than the metallization width, and the heat flow becomes dominated by the silicon film. While the model developed by Goodson and Flik allows separate input parameters for each width, it does not accurately model heat flow for the case where the device width is greater than the width of the metallization. The first part of this chapter re-examines the Goodson and Flik model and presents a modification to correct for the case where the device width is greater than the width of the drain/ source interconnects. The impact of the modification is examined through comparisons to the original model and measurements. The second part of the chapter then details the derivation of the dynamic thermal impedance model. The accuracy of the dynamic model is supported with three-dimensional finite-element simulations and measurements.

6.2   Derivation of the SOI MOSFET Thermal Resistance Model

The thermal resistance model represents a solitary SOI MOSFET with three parallel conductance paths that carry heat away from the channel. These heat-flow paths are illustrated in Figure 6.2 and are modeled by a system of coupled one-dimensional differential equations. The derivation of the thermal resistance model is divided into separate regions that correspond to the different thermal conductance

Channel temperature, T_c

Along drain, source and interconnects

Thru buried oxide beneath channel

Along gate interconnect

Substrate temperature, T_0
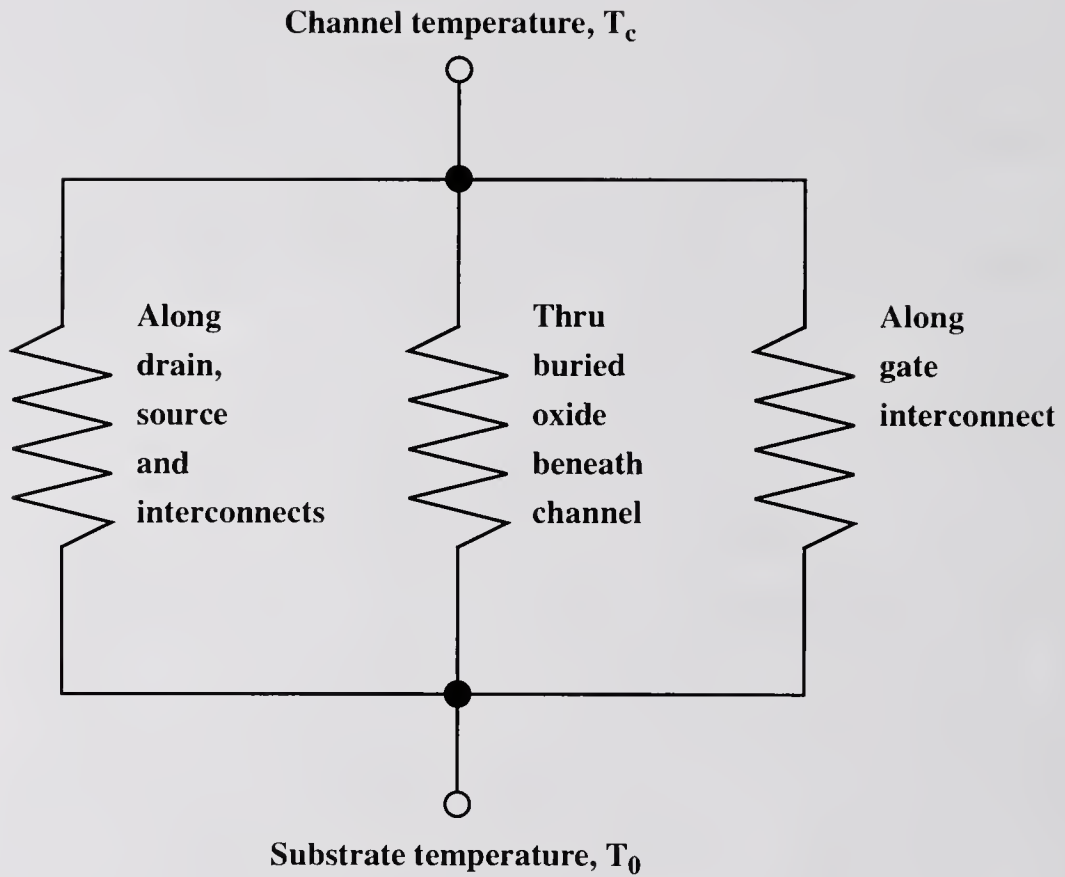
Figure 6.2   Illustration of the three parallel conductance paths for heat to travel from the channel to the substrate

paths. Different subscripts are used to denote the equations and variables that apply to the different regions; the subscript f is used as a generic identifier, and the subscripts c, g, d, m and ct apply to the channel, gate interconnect, drain/source regions of the silicon film, drain/source interconnects and the drain/source contact regions, respectively. Figure 6.3 shows the simplified device structure that defines the geometry for the model derivation. The definitions of the model parameters are listed in Table 6.1.

The channel region is modeled as a heat source with uniform power dissipation and is assumed to be at a uniform temperature T_c. In reality, the heat generation rate in the channel has a strong spatial dependence due to variations in the surface potential along the channel. However, the isothermal channel approximation was shown to be valid for predicting the average channel temperature, giving results typically within 10% error [Goo95]. The interface between the buried oxide and the silicon substrate is assumed to be at a uniform temperature T_0, and the device is assumed to cool solely through the substrate. As shown in Chapter Two, the thermal conductance to the overlying oxide layers is generally much smaller than the conductance to the substrate; therefore, heat loss through the overlying oxide layers can be neglected.

The gate interconnect, the drain and source, and their metal interconnects are treated as one-dimensional cooling fins that carry heat away from the channel. The differential equations that describe the temperature rise along the respective fins are given by
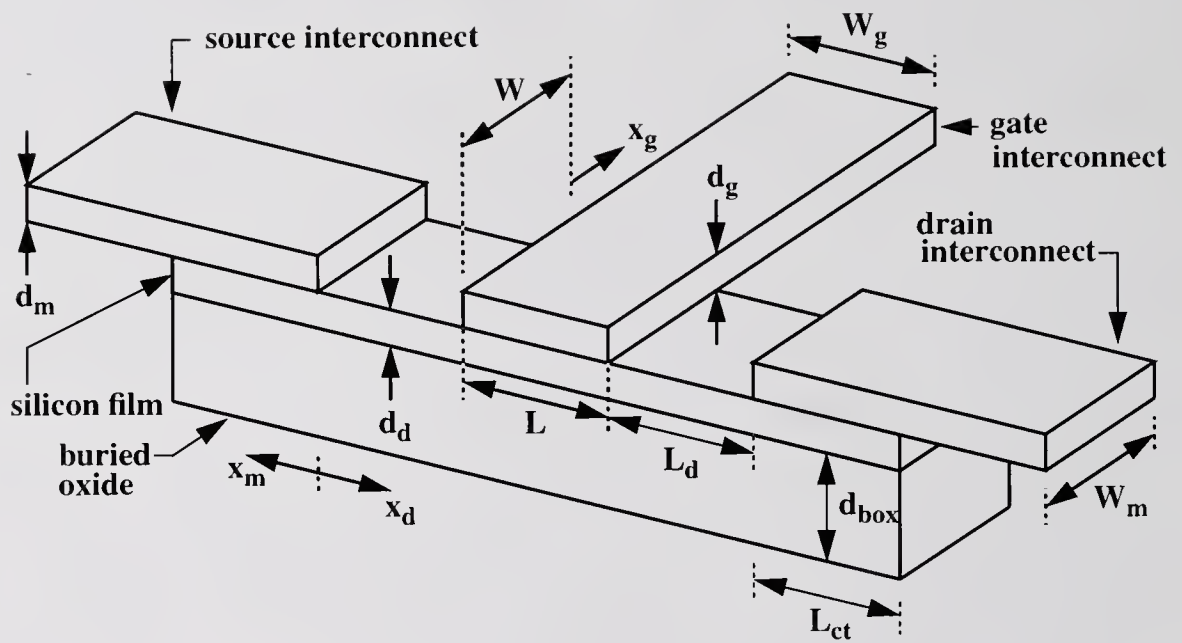
Figure 6.3   The simplified device geometry used to define the solution domain for the SOI MOSFET thermal resistance model. The model parameters are defined in Table 6.1.

Table 6.1 SOI MOSFET thermal resistance model parameters

Parameter	Definition
W	Width of channel
W_m	Width of drain/source interconnect
W_g	Width of gate interconnect
L	Length of channel
L_d	Length between gate and drain/source contact
L_{ct}	Length of drain/source contact opening
d_{box}	Thickness of buried oxide
d_{gox}	Thickness of gate oxide
d_{fox}	Thickness of field oxide
d_{LTO}	Thickness of oxide between fin and field oxide
d_d	Thickness of silicon film
d_g	Thickness of gate interconnect
d_m	Thickness of drain/source interconnect

$$\frac{\partial^2 \Delta T_d(x_d)}{\partial x_d^2} - m_d^2 \Delta T_d(x_d) = 0, \tag{6.1}$$

$$\frac{\partial^2 \Delta T_m(x_m)}{\partial x_m^2} - m_m^2 \Delta T_m(x_m) = 0 \tag{6.2}$$

and

$$\frac{\partial^2 \Delta T_g(x_g)}{\partial x_g^2} - m_g^2 \Delta T_g(x_g) = 0, \tag{6.3}$$

where $\Delta T_f(x_f) = T_f(x_f) - T_0$. Since the model represents an isolated device, the drain/source and gate interconnects are assumed to be infinitely long, so that

$$\Delta T_g(x_g)\big|_{x_g \to \infty} = 0 \tag{6.4}$$

and

$$\Delta T_m(x_m)\big|_{x_m \to \infty} = 0. \tag{6.5}$$

The second term on the left-hand side of (6.1), (6.2) and (6.3), accounts for heat conduction through the underlying oxide as the heat travels along the fin, where

$$\frac{1}{m_f} = \sqrt{\frac{k_f d_f}{h_f}} \tag{6.6}$$

is the characteristic thermal length of the fin and

$$h_f = \frac{k_o}{d_{fo}} \tag{6.7}$$

is the heat transfer coefficient from the fin to the substrate. The heat transfer coefficient is derived by solving the one-dimensional, steady-state heat conduction equation in the underlying oxide layers. Values for the necessary material properties are listed in Table 6.2. The variable d_{fo} corresponds to the total oxide thickness between a fin and the substrate, so that

$$d_{do} = d_{box}, \tag{6.8}$$

$$d_{mo} = d_{box} + d_{fox} + d_{LTO}, \tag{6.9}$$

and

$$d_{go} = d_{box} + d_{fox}. \tag{6.10}$$

Assuming that the temperature along each fin is one-dimensional, neglects the temperature gradients in the vertical and lateral directions within each fin. The validity of this assumption can be evaluated using the Biot number, which corresponds to the ratio of the internal and external thermal resistances for a given fin [Ozi93]. If the Biot number is much less than unity, then the fin can be approximated by a one-dimensional thermal medium. The vertical Biot number for a fin is expressed as $B_{Vf} = h_f d_f / k_f$. For B_{Vf} to equal 0.1, the thickness of the fin (d_f) would have to be approximately 5, 12 and 2 times greater than d_{fo} for the drain/source, drain/source interconnects and the gate interconnect, respectively.

Table 6.2 Material properties

Property	Definition	Value
k_o	Thermal conductivity of silicon dioxide	0.014 (W cm^{-1} K^{-1})
k_d	Thermal conductivity of drain/source silicon film	0.63 (W cm^{-1} K^{-1})
k_m	Thermal conductivity of drain/source interconnect material[‡]	2.39 (W cm^{-1} K^{-1})
k_g	Thermal conductivity of gate interconnect material[‡‡]	0.30 (W cm^{-1} K^{-1})

Source: [Goo95]

[‡] Assumed to be aluminum

[‡‡] Assumed to be doped polysilicon

Consequently, for typical SOI MOSFET geometries, the vertical temperature gradients in the fins can be neglected. The lateral Biot number for a fin is given by $B_{Lf} = h_f W_f^2/(k_f d_f)$. For the drain/source interconnects, assuming they are routed in the first metal layer, B_{Lm} approaches unity when $W_m \cong 13\sqrt{d_m}$; for practical metallization geometries, the lateral temperature gradient can be neglected. For a highly-doped polysilicon gate interconnect, the relatively low thermal conductivity can lead to more significant lateral temperature gradients. However, these gradients should be substantial only for channel lengths greater than 1.2 μm. Due to the effects of the drain/source contacts and interconnects, the heat flow in the silicon film is more complex and the lateral Biot number is not valid for these regions. ANSYS simulations of the drain/source silicon film regions are used to evaluate the nature of the heat flow; the results are shown in Figure 6.4. The simulations show that for low thermal conductances to the substrate (i.e thicker buried oxides), the effect of the drain/source interconnects can lead to significant lateral temperature gradients in the silicon film. However, the model accurately predicts the average temperature distribution in the drain/source silicon film, showing that the one-dimensional approximation is reasonable for these regions.

The temperature rise of the silicon film at the edge of the channel ($x_d = L_d$) and of the gate interconnect at $x_g = 0$ are given by

$$\Delta T_d(x_d)\Big|_{x_d = L_d} = \Delta T_g(x_g) \cdot \left[\frac{1}{1 + \frac{W_g d_{gox}}{k_{gox} W \cdot L}\sqrt{\frac{d_g k_g k_o}{d_{go}}}} \right]^{-1}\Bigg|_{x_g = 0} = \Delta T_c. \qquad (6.11)$$
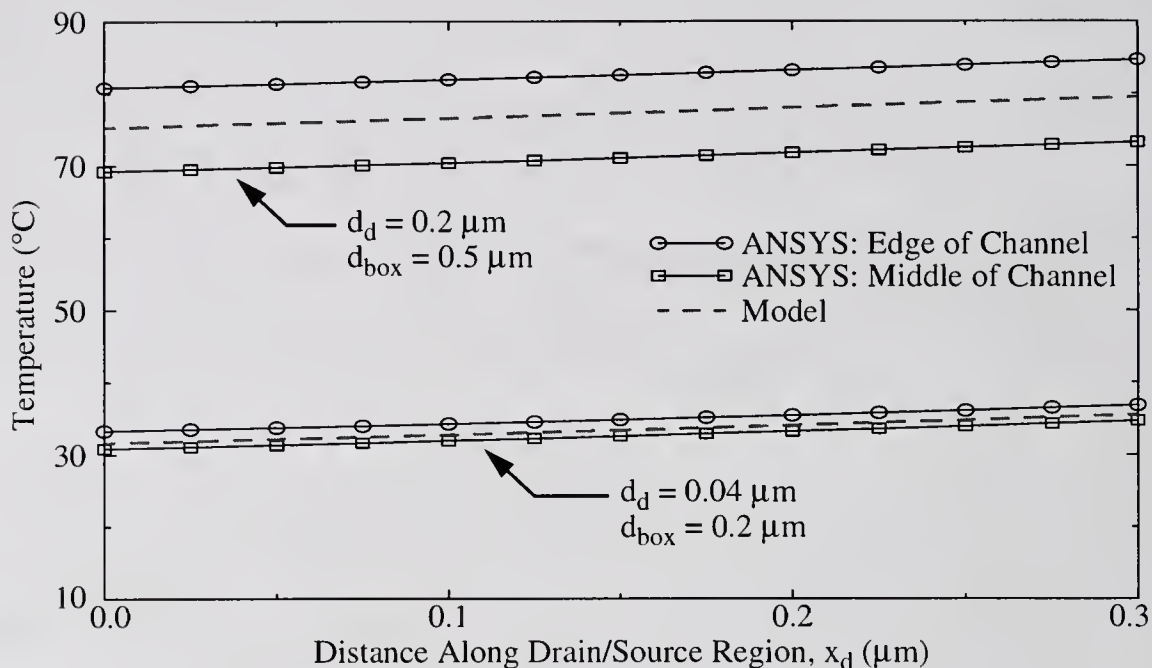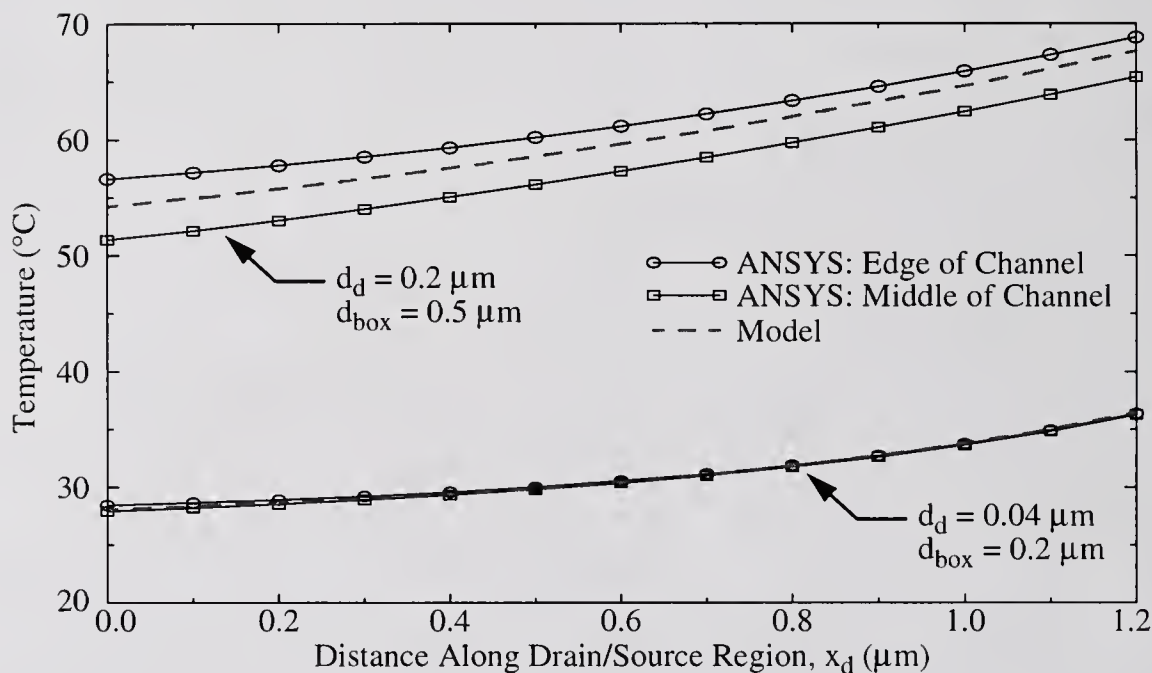
Figure 6.4   Temperature along drain/source regions as calculated by ANSYS and predicted by the model for an SOI MOSFET with W = 12 μm and W_m = 2 μm (center contact). The ANSYS data is evaluated at both the center and the edge of the channel, to show the extent of the lateral temperature gradient: a) For L_d = 0.3 μm; b) for L_d = 1.2 μm.

where k_{gox} is the thermal conductivity of the gate oxide. The multiplier for the gate interconnect temperature accounts for the thermal resistance of the gate oxide by assuming

$$\frac{\partial^2 \Delta T_{gox}}{\partial x_{gox}^2} = 0, \tag{6.12}$$

$$\Delta T_{gox}\Big|_{x_{gox} = 0} = \Delta T_c \tag{6.13}$$

and

$$\Delta T_{gox}\Big|_{x_{gox} = d_{gox}} = \Delta T_g(x_g)\Big|_{x_g = 0}. \tag{6.14}$$

Goodson and Flik assumed that the thermal resistance of the gate oxide was negligible since d_{gox} is typically very small [Goo92]. However, (6.11) shows that the effect of the gate oxide thermal resistance can become more significant if k_{gox} decreases or k_g increases; both conditions are probable in advanced SOI MOSFET's due to phonon boundary scattering in the gate oxide and the use of silicides for the gate interconnect material. The effect of the gate oxide can be removed by setting $d_{gox} = 0$, in which case, the model reverts to the original model developed by Goodson and Flik [Goo92].

At the interface between the silicon film and the drain/source interconnect metallization ($x_d = x_m = 0$), the temperatures of the silicon and the metal are assumed to be equal and the heat flows are equated, giving the boundary conditions

$$\Delta T_m(x_m)\big|_{x_m = 0} = \Delta T_d(x_d)\big|_{x_d = 0} \qquad (6.15)$$

$$-k_m W_m d_m \left[\frac{\partial \Delta T_m(x_m)}{\partial x_m}\right]\Bigg|_{x_m = 0} - k_d W d_d \left[\frac{\partial \Delta T_d(x_d)}{\partial x_d}\right]\Bigg|_{x_d = 0} = 0. \qquad (6.16)$$

Equation (6.16) is the boundary condition derived by Goodson and Flik, and is valid when $W \approx W_m$ [Goo92]. However, since the thermal analysis for the drain/source silicon film region only extends to the edge of the interconnect contact, when $W > W_m$ (6.16) neglects heat flow from the portions of the silicon film under the contact. Figure 6.5 illustrates the portions of the silicon film that are neglected by (6.16). The heat flow from these portions of the silicon film through the buried oxide is proportional to $W - W_m$; thus, the error introduced by neglecting this heat flow increases as W becomes greater than W_m.

To account for heat flow from the portions of the drain/source silicon film under the contact, (6.16) is modified. Due to the high thermal conductivity of the metal contact, the silicon film under the contact is assumed to be at a uniform temperature. For large devices where the contact area is divided into multiple contact windows, significant lateral temperature gradients can develop. ANSYS simulations are used to examine the heat flow in the silicon region under the contact; the results are shown in Figure 6.6. Even for conditions that cause considerable lateral heat flow, the model gives an accurate account of the average temperature in the silicon film under the contact. Therefore, the uniform temperature assumption can be considered appropriate. To account for the neglected conduction, an additional heat-flow term is added to (6.16), so that the modified boundary condition becomes
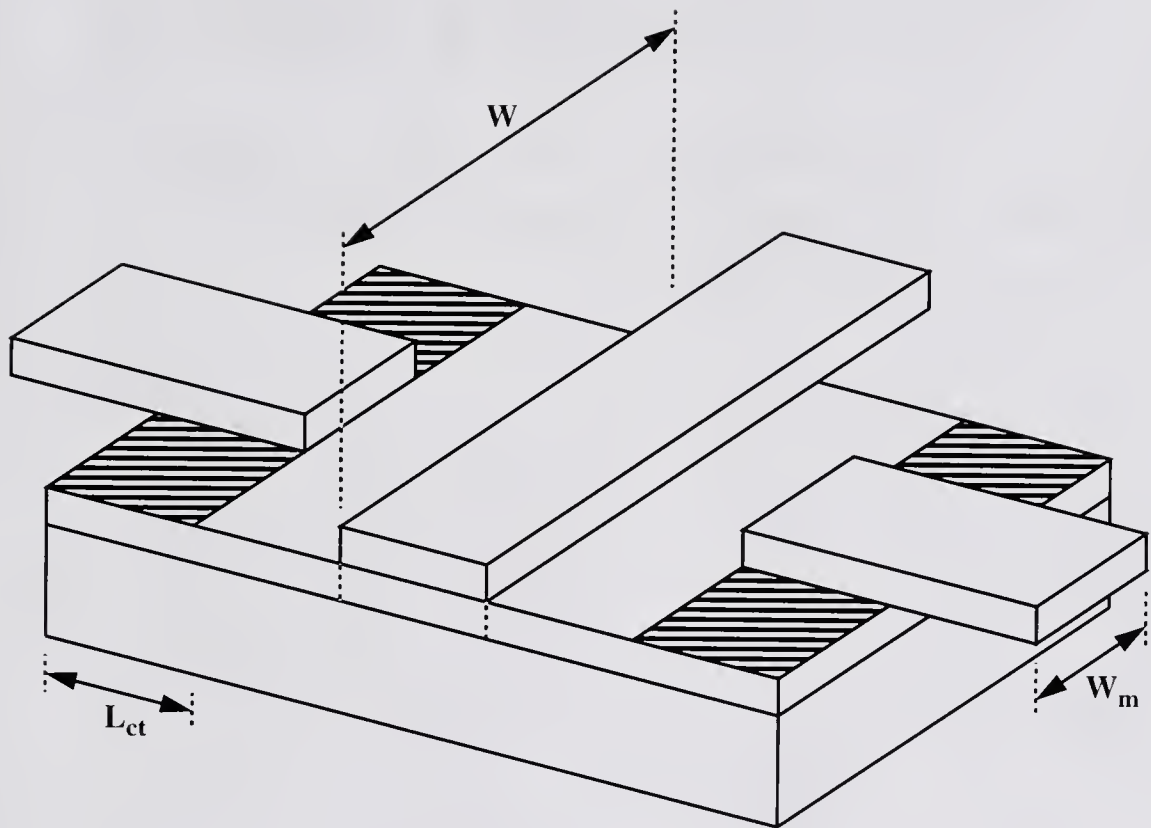
Figure 6.5   The simplified SOI MOSFET structure for the case where $W > W_m$. The cross-hatched regions represent the portions of the drain/source silicon film where heat flow is neglected by the Goodson and Flik model [Goo92].
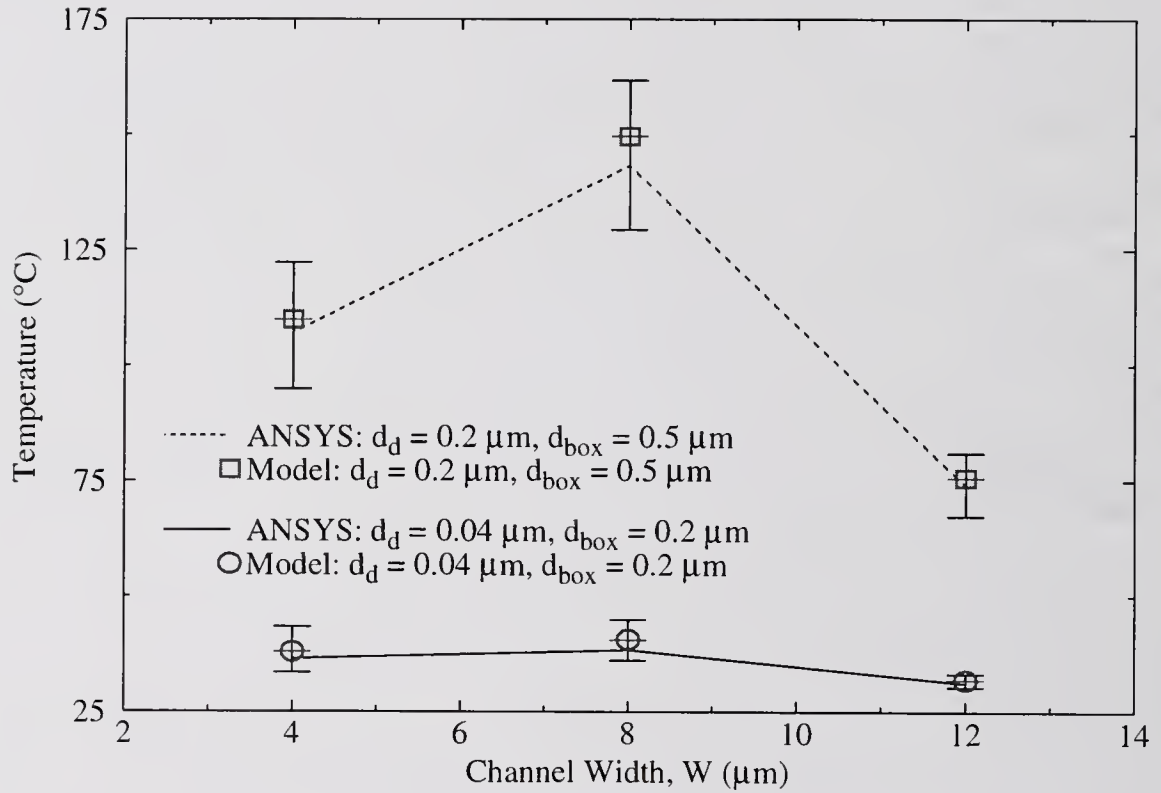
Figure 6.6   The average temperature of the silicon film under the drain/source contacts as calculated by ANSYS and predicted by the model for an SOI MOSFET with $L_d = 0.3$ μm and $L_{ct} = 0.6$ μm. The error bars show the highest and lowest temperatures calculated by ANSYS for these regions.

$$0 = -k_m W_m d_m \left[\frac{\partial \Delta T_m(x_m)}{\partial x_m}\right]\Bigg|_{x_m = 0} - k_d W d_d \left[\frac{\partial \Delta T_d(x_d)}{\partial x_d}\right]\Bigg|_{x_d = 0}. \qquad (6.17)$$
$$+ h_d(W - W_m)L_{ct}\Delta T_d(x_d)\Big|_{x_d = 0}$$

For $W \leq W_m$, the last term on the right-hand side of (6.17) is forced to reduce to zero such that the model reverts to the original boundary condition developed by Goodson and Flik [Goo92].

The boundary conditions (6.4), (6.5), (6.11), (6.15) and (6.17), are used to solve (6.1), (6.2) and (6.3), giving the temperature rise along each cooling fin. The thermal conductance paths are then coupled through the following power conservation equation

$$2k_d W d_d \left[\frac{\partial \Delta T_d(x_d)}{\partial x_d}\right]\Bigg|_{x_d = L_d} - k_g W_g d_g \left[\frac{\partial \Delta T_g(x_g)}{\partial x_g}\right]\Bigg|_{x_g = 0} + h_d W \cdot L \Delta T_c = P. \quad (6.18)$$

Equation (6.18) can be solved in the form

$$\Delta T_c = R_{TH} P. \qquad (6.19)$$

The resulting expression for the thermal resistance is given by

$$R_{TH} = \left[\frac{2m_d k_d d_d W(\arg 1 - \arg 2)}{\arg 1 + \arg 2} + \frac{k_{gox} W \cdot L}{d_{gox} + \dfrac{k_{gox} W \cdot L}{k_g m_g W_g d_g}} + h_d W \cdot L\right]^{-1} \qquad (6.20)$$

where

$$\arg 1 = \exp(m_d L_d)(m_d k_d d_d W + h_d(W - W_m)L_{ct} + m_m k_m d_m W_m) \qquad (6.21)$$

and

$$\arg 2 \; = \; \exp(-m_d L_d)(m_d k_d d_d W - h_d (W - W_m) L_{ct} - m_m k_m d_m W_m). \tag{6.22}$$

For the case where $W \le W_m$, the expression for the thermal resistance is equivalent to the Goodson and Flik model [Goo92] and

$$\arg 1 \; = \; \exp(m_d L_d)(m_d k_d d_d W + m_m k_m d_m W_m) \tag{6.23}$$

$$\arg 2 \; = \; \exp(-m_d L_d)(m_d k_d d_d W - m_m k_m d_m W_m). \tag{6.24}$$

The differential equations that describe the temperature rise in each cooling fin neglect the temperature dependences of the material thermal conductivities. However, the measured data plotted in Figure 6.7 show that the temperature dependence of the SOI MOSFET thermal resistance is relatively weak. For the devices used in the measurements, the extracted thermal resistances vary by less than 3% over a 90 °C temperature range. Therefore, it is assumed that the temperature dependences of the thermal conductivities can be neglected.

6.3 Verification of the SOI MOSFET Thermal Resistance Model

The accuracy of the SOI MOSFET thermal resistance model derived by Goodson and Flik has been verified [Goo92, Goo95]. The model was shown to agree with thermal resistance measurements of devices with $W \le W_m$. To validate the modification described in the previous section, the new model is compared to extracted thermal resistances of large devices. The thermal resistance of SOI
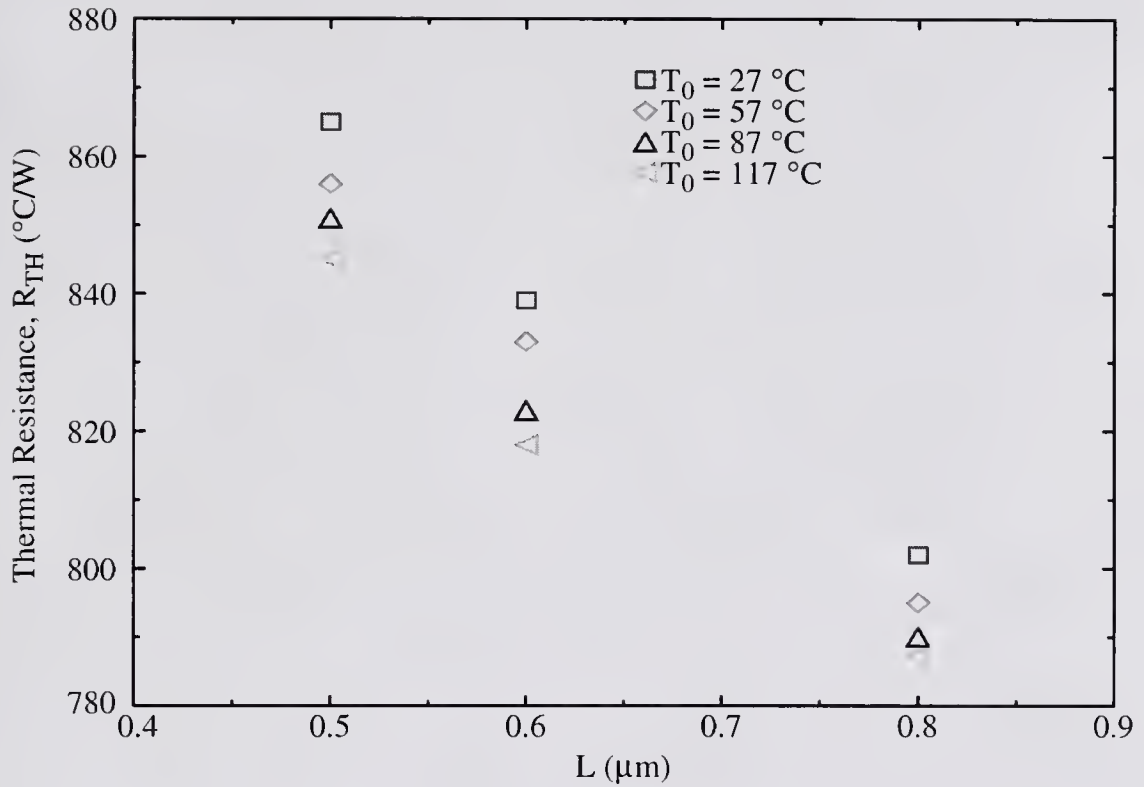
Figure 6.7   The temperature dependence of the SOI MOSFET thermal resistance. The data was extracted using the gate resistance thermometry technique [Goo95]. The device parameters were W = 120 μm, d_{box} = 0.4 μm and d_d = 0.1 μm.

MOSFET's with $W = 120 \ \mu m$ and $W_m = 2 \ \mu m$ were measured using a four-point, gate resistance thermometry technique [Goo95]. Figure 6.8 compares both models to the measured thermal resistance values. By neglecting the heat flow out of the contact regions, the model developed by Goodson and Flik [Goo92] tends to over-estimate the thermal resistance, in this case by as much as 40%. With modification to account for the heat flow out of the drain/source contact regions implemented, the model provides a much more accurate estimation of the thermal resistance. For the given devices, the error is reduced to no more than three percent.

6.4 Derivation of the SOI MOSFET Thermal Impedance Model

The thermal impedance model for SOI MOSFET's is derived by extending the steady-state thermal resistance model into the time domain. The time-dependent form of the differential cooling-fin equation can be expressed as

$$\frac{\partial^2 \Delta T_f}{\partial x_f^2} - [m_f(t)]^2 \Delta T_f = \frac{1}{\alpha_f} \frac{\partial \Delta T_f}{\partial t} \qquad (6.25)$$

where $\alpha_f = k_f/(\rho_f c_{pf})$ is the thermal diffusivity of the fin material. Values for the additional material properties needed to calculate the thermal diffusivities are listed in Table 6.3. In the time domain, (6.25) does not have a simple closed-form solution. However, a simple solution to (6.25) can be found in the frequency domain using the Laplace transform.
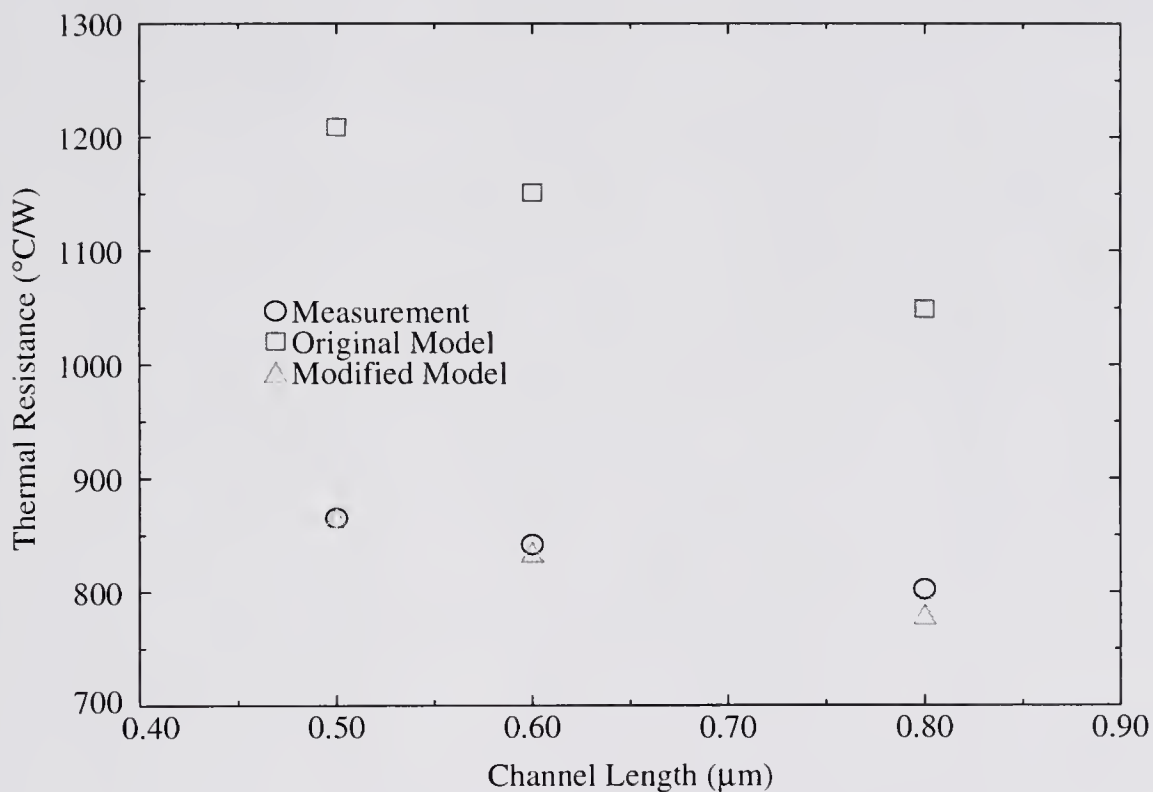
Figure 6.8   Measured and simulated data for the thermal resistance of Motorola SOI MOSFET's. The modified model is compared to the original Goodson and Flik model [Goo92]. The measured thermal resistance values were extracted using the gate resistance thermometry technique [Goo95]. The device parameters were $W = 120 \, \mu m$, $W_m = 2 \, \mu m$, $d_{box} = 0.4 \, \mu m$ and $d_d = 0.1 \, \mu m$.

Table 6.3 Material properties

Property	Definition	Value
ρ_o	Density of silicon dioxide	2.19 (g cm^{-3})
ρ_d	Density of silicon film	2.328 (g cm^{-3})
ρ_m	Density of drain/source interconnect material[‡‡]	2.7[‡] (g cm^{-3})
ρ_g	Density of gate interconnect material[‡‡‡]	2.328 (g cm^{-3})
c_{po}	Specific heat of silicon dioxide	1.4 (J g^{-1} K^{-1})
c_{pd}	Specific heat of silicon film	0.7 (J g^{-1} K^{-1})
c_{pm}	Specific heat of drain/source interconnect material[‡‡]	0.9[‡] (J g^{-1} K^{-1})
c_{pg}	Specific heat of gate interconnect material[‡‡‡]	0.7 (J g^{-1} K^{-1})

Source: [Mul77] (unless specified otherwise)

[‡] [Ozi93]

[‡‡] Assumed to be aluminum

[‡‡‡] Assumed to be doped polysilicon

In the frequency domain, the differential equations (6.1), (6.2) and (6.3) become

$$\frac{\partial^2 \theta_d(x_d, s)}{\partial x_d^2} - \left\{ [m_d(s)]^2 + \frac{s}{\alpha_d} \right\} \theta_d(x_d, s) = 0 \tag{6.26}$$

$$\frac{\partial^2 \theta_m(x_m, s)}{\partial x_m^2} - \left\{ [m_m(s)]^2 + \frac{s}{\alpha_m} \right\} \theta_m(x_m, s) = 0 \tag{6.27}$$

$$\frac{\partial^2 \theta_g(x_g, s)}{\partial x_g^2} - \left\{ [m_g(s)]^2 + \frac{s}{\alpha_g} \right\} \theta_g(x_g, s) = 0 \tag{6.28}$$

where $\theta_f(x_f, s) = \pounds[\Delta T_f(x_f, t)]$ is the Laplace transform of the time-domain temperature rise and the initial condition in each fin is specified as $\Delta T_f(x_f, 0) = 0$ [Ozi93]. Examining (6.26), (6.27) and (6.28), shows that the differential cooling fin equation can be expressed in the same form for either the steady-state or the frequency domain. Therefore, the frequency-domain differential cooling-fin equations have closed-form general solutions given as

$$\theta_f(x_f, s) = c_1 \exp(\gamma_f x_f) + c_2 \exp(-\gamma_f x_f) \tag{6.29}$$

where

$$\gamma_f = \sqrt{[m_f(s)]^2 + \frac{s}{\alpha_f}} \tag{6.30}$$

and c_1 and c_2 are arbitrary constants that depend on the boundary conditions. The use of the Laplace transform requires that the temperature rise remains a linear function of the power dissipation over the desired region of operation. Since the thermal resistance only depends weakly on device bias and temperature, this assumption is valid for typical device operation.

As mentioned in the first part of this chapter, the characteristic thermal length, $1/m_f$, of a cooling fin accounts for heat conduction through the underlying oxide layers. During a thermal transient, the characteristic thermal length will vary, $1/m_f = 1/m_f(s)$, due to the time dependence of the heat transfer in the underlying oxide. Using the steady-state heat transfer coefficient to calculate the characteristic thermal length would imply that the heat flux in the buried oxide instantaneously adjusts to a steady-state value. Therefore, the heat transfer coefficient for the fin, $h_f(s)$, becomes a function of frequency and should be derived from a dynamic analysis of the heat flux in the underlying oxide.

Assuming the heat flow in the oxide is predominantly one-dimensional, the temperature rise in the underlying oxide is described by

$$\frac{\partial^2 \Delta T_o(n, t)}{\partial n^2} = \frac{1}{\alpha_o} \frac{\partial \Delta T_o(n, t)}{\partial t} \tag{6.31}$$

where $\Delta T_o(n, t) = T_o(n, t) - T_0$, α_o is the thermal diffusivity of silicon dioxide and n is the direction of heat flow. Taking the Laplace transform of (6.31) results in

$$\frac{\partial^2 \theta_o(n, s)}{\partial n^2} - \frac{s}{\alpha_o} \theta_o(n, s) = 0, \tag{6.32}$$

where the initial temperature rise in the oxide is assumed to be $\Delta T_o(n, 0) = 0$. The fin and the underlying oxide are assumed to be in perfect thermal contact, so that at the interface between the fin and the oxide

$$\theta_o(s) = \theta_f(s) . \tag{6.33}$$

The interface between the buried oxide and the substrate is still assumed to be at the temperature T_0, so at that interface

$$\theta_o(s) = 0 . \tag{6.34}$$

The heat flux at the interface between the fin and the oxide can be derived from the solution to (6.32) for the given boundary conditions, and is given by

$$-k_o \frac{\partial \theta_o(n, s)}{\partial n} = h_f(s) \cdot \theta_f(s) = k_o \sqrt{\frac{s}{\alpha_o}} \left[\frac{1 + \exp\left(-2\sqrt{\frac{s}{\alpha_o}} d_{fo}\right)}{1 - \exp\left(-2\sqrt{\frac{s}{\alpha_o}} d_{fo}\right)} \right] \theta_f(s) . \tag{6.35}$$

Therefore, the dynamic heat transfer coefficient can be expressed as

$$h_f(s) = k_o \sqrt{\frac{s}{\alpha_o}} \coth\left(d_{fo} \sqrt{\frac{s}{\alpha_o}}\right) . \tag{6.36}$$

The gate interconnect and the drain/source interconnects are still assumed to be infinitely long, so that

$$\theta_g(x_g, s)\big|_{x_g \to \infty} = 0 \tag{6.37}$$

and

$$\theta_m(x_d, s)\big|_{x_m \to \infty} = 0. \qquad (6.38)$$

The boundary condition at the edge of the channel becomes

$$\theta_d(x_d, s)\big|_{x_d = L_d} = \Gamma_{gox1} \cdot \theta_g(x_g, s)\big|_{x_g = 0} = \theta_c(s) \qquad (6.39)$$

where

$$\Gamma_{gox1} = \frac{1 + \exp(-2\gamma_{gox}d_{gox})}{2\exp(-\gamma_{gox}d_{gox})} + \frac{k_g\gamma_g W_g d_g[1 - \exp(-2\gamma_{gox}d_{gox})]}{2k_{gox}\gamma_{gox}W \cdot L\exp(-\gamma_{gox}d_{gox})} \qquad (6.40)$$

and

$$\gamma_{gox} = \sqrt{\frac{s}{\alpha_{gox}}}. \qquad (6.41)$$

Equation (6.40) is derived by solving (6.32) in the gate oxide for the boundary conditions implied by (6.13) and (6.14). The boundary conditions that define the temperature and heat flow at the interface between the drain/source regions and the interconnect metallization become

$$\theta_m(x_m, s)\big|_{x_m = 0} = \theta_d(x_d, s)\big|_{x_d = 0} \qquad (6.42)$$

and

$$0 = -k_m W_m d_m \left[\frac{\partial \theta_m(x_m, s)}{\partial x_m}\right]\Bigg|_{x_m = 0} - k_d W d_d \left[\frac{\partial \theta_d(x_d, s)}{\partial x_d}\right]\Bigg|_{x_d = 0}.$$
$$+ h_d(s)[W - W_m]L_{ct}\theta_d(x_d, s)\Big|_{x_d = 0} \tag{6.43}$$

The Laplace transform of the power conservation equation that couples the conductance paths is

$$p(s) = 2k_d W d_d \left[\frac{\partial \theta_d(x_d, s)}{\partial x_d}\right]\Bigg|_{x_d = L_d} - k_g W_g d_g \left[\frac{\partial \theta_g(x_g, s)}{\partial x_g}\right] \cdot \Gamma_{gox2}\Bigg|_{x_g = 0} \tag{6.44}$$
$$+ h_d(s)WL\theta_c(s)$$

where

$$\Gamma_{gox2} = \frac{2\exp(-\gamma_{gox}d_{gox})}{[\exp(-2\gamma_{gox}d_{gox}) + 1] - \dfrac{k_{gox}\gamma_{gox}W \cdot L}{k_g\gamma_g d_g W_g}[\exp(-2\gamma_{gox}d_{gox}) - 1]} \tag{6.45}$$

and $p(s) = £[p(t)]$ is the Laplace transform of the instantaneous power dissipation in the channel. Equation (6.44) can be solved in the form

$$\Delta T_c = Z_{TH}(s) \cdot p(s), \tag{6.46}$$

resulting in the following expression for the dynamic thermal impedance

$$Z_{TH}(s) = \left[\frac{2(\arg 1 - \arg 2)k_d\gamma_d d_d W}{\arg 1 + \arg 2} + k_g\gamma_g d_g W_g\Gamma_{gox2} + h_d(s)W \cdot L\right]^{-1} \tag{6.47}$$

where

$$\text{arg}1 = \exp[L_d\gamma_d]\{k_d\gamma_d d_d W + h_d(s)[W - W_m]L_{ct} + k_m\gamma_m d_m W_m\} \tag{6.48}$$

and

$$\text{arg}2 = \exp[-L_d\gamma_d]\{k_d\gamma_d d_d W - h_d(s)[W - W_m]L_{ct} - k_m\gamma_m d_m W_m\}. \tag{6.49}$$

6.5 Verification of the SOI MOSFET Thermal Impedance Model

To verify the thermal impedance model, three-dimensional (3-D) finite-element (FE) simulations of an SOI MOSFET were performed using ANSYS. The FE model represented a simplified device structure and was constructed using similar assumptions and boundary conditions to those detailed in Chapter Two. The model was further simplified by neglecting the gate interconnect. Figure 6.9a shows the FE model used to represent the SOI MOSFET structure in ANSYS. The thermal impedance was extracted from the FE simulations by averaging the temperature rise over the channel region. The ANSYS results are compared to (6.47) with the term corresponding to conduction in the gate interconnect set to zero. The comparisons are illustrated in Figure 6.9b, Figure 6.10, and Figure 6.11. The SOI MOSFET thermal impedance model displays reasonable trends with the scaling of the silicon film thickness, the buried oxide thickness, and both channel width and length. The model does a good job of reproducing both the transient and steady-state temperature responses produced by the 3-D FE simulations. For the device structures that were simulated, the error between the model and ANSYS did not exceed 14%. As is evident from the comparisons in Figure 6.10, the discrepancy between the model and the FE simulations is greatest during the latter half of the transient. This trend can be
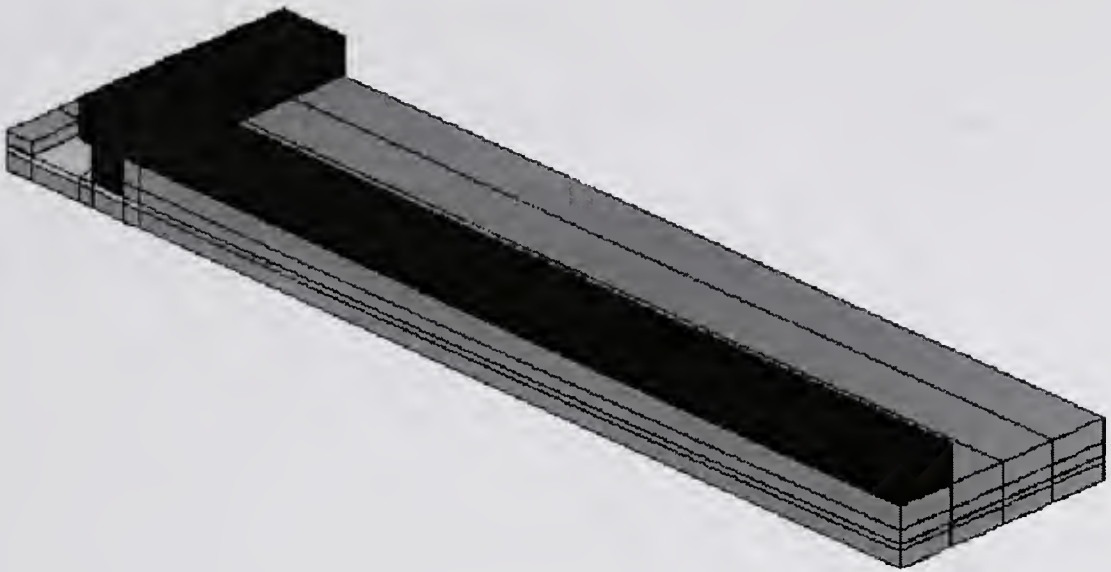
(a)
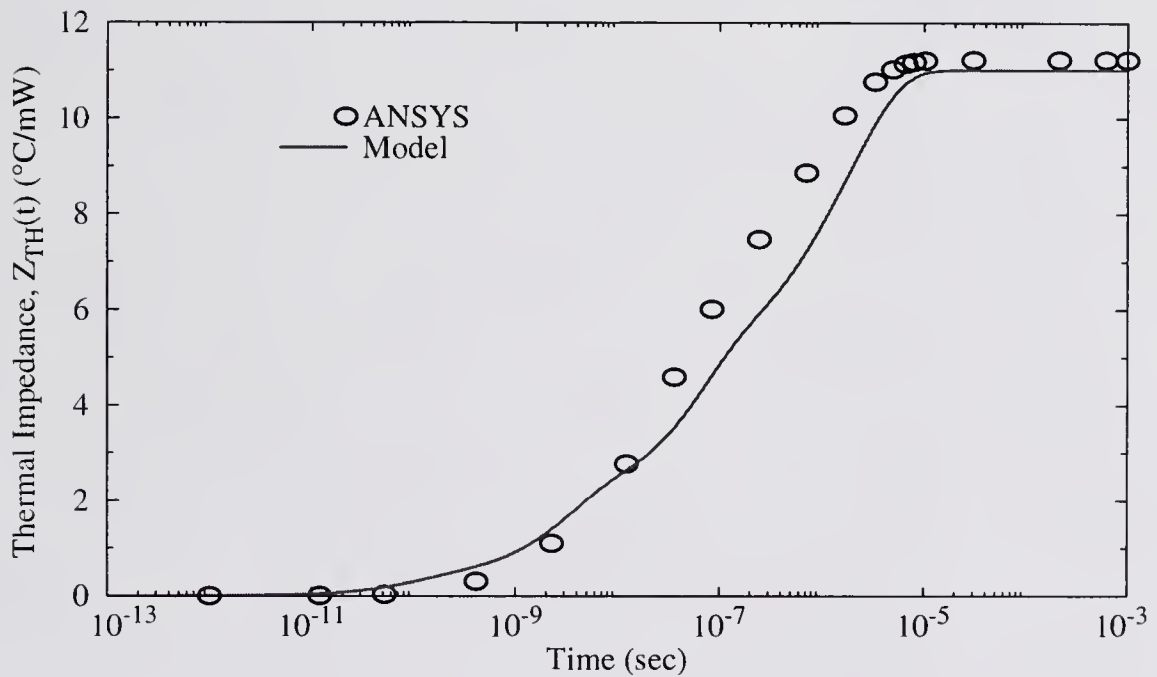
(b)

Figure 6.9   The transient thermal impedance simulated with ANSYS and calculated with the SOI MOSFET thermal impedance model: a) The finite-element model simulated with ANSYS; b) comparison of the thermal impedance simulated with ANSYS and the model for $W = 8$ μm, $L = 0.35$ μm, $L_d = 0.6$ μm, $d_d = 0.15$ μm, $d_{box} = 0.4$ μm, $W_m = 2$ μm and $d_m = 0.9$ μm.
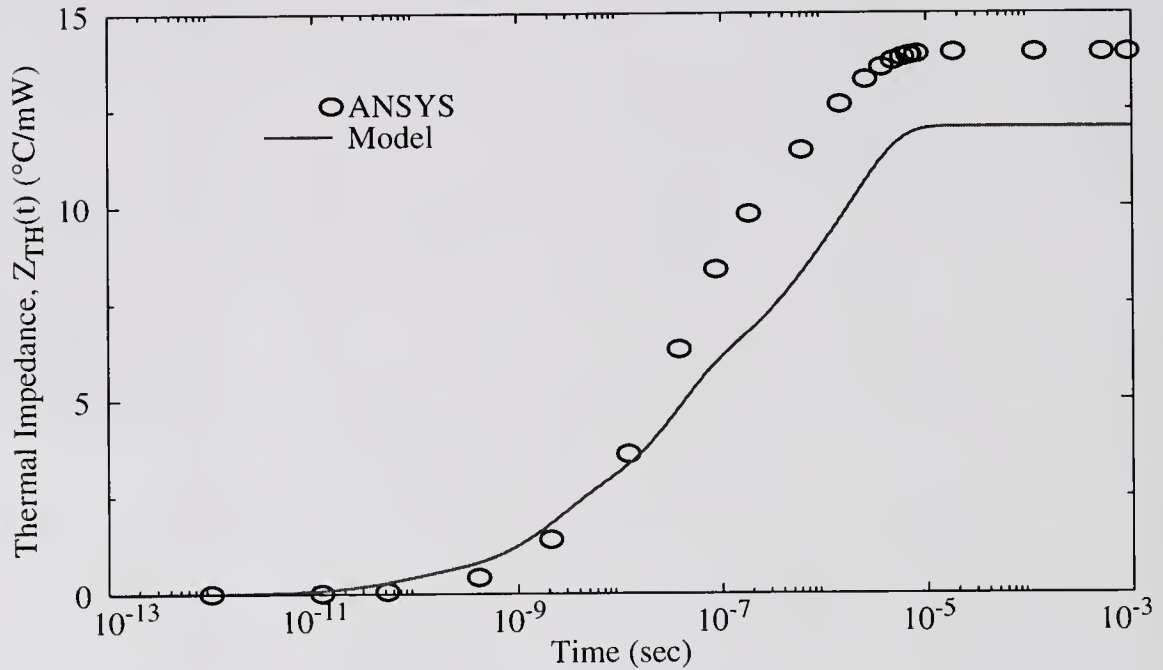
(a)

(b)

Figure 6.10   The transient thermal impedance simulated with ANSYS and calculated with the SOI MOSFET thermal impedance model for $W = 8$ μm, $L = 0.35$ μm, $L_d = 0.6$ μm, $d_d = 0.1$ μm, $W_m = 2$ μm and $d_m = 0.9$ μm: a) For $d_{box} = 0.4$ μm; b) for $d_{box} = 0.2$ μm.
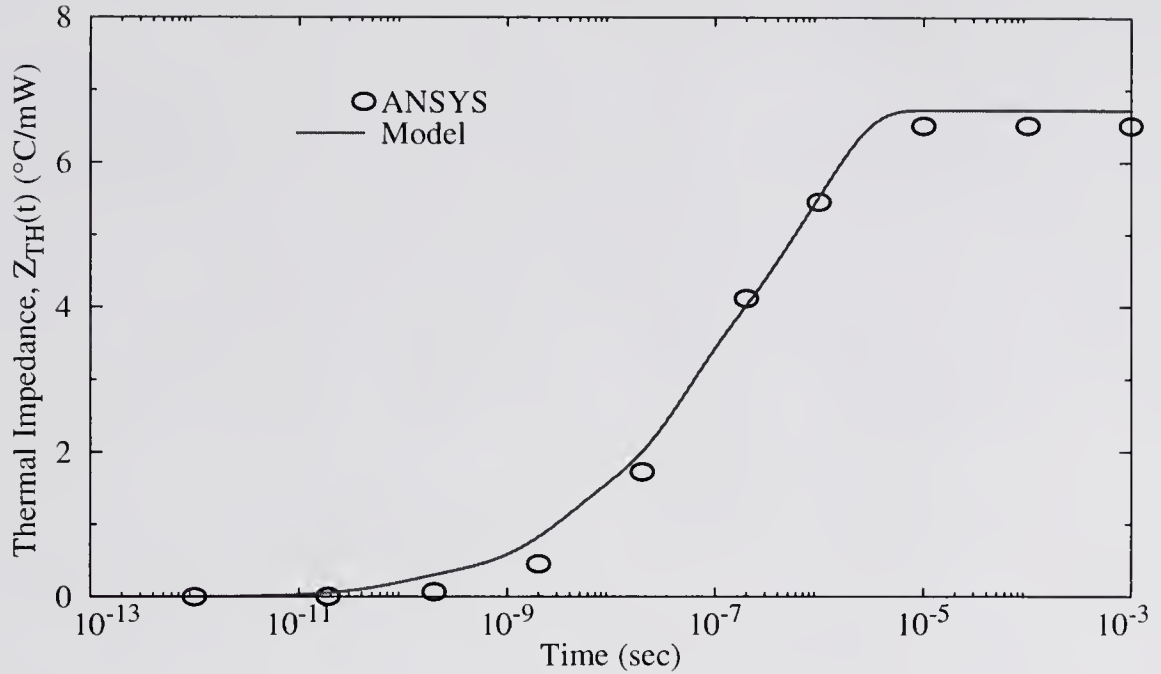
(a)
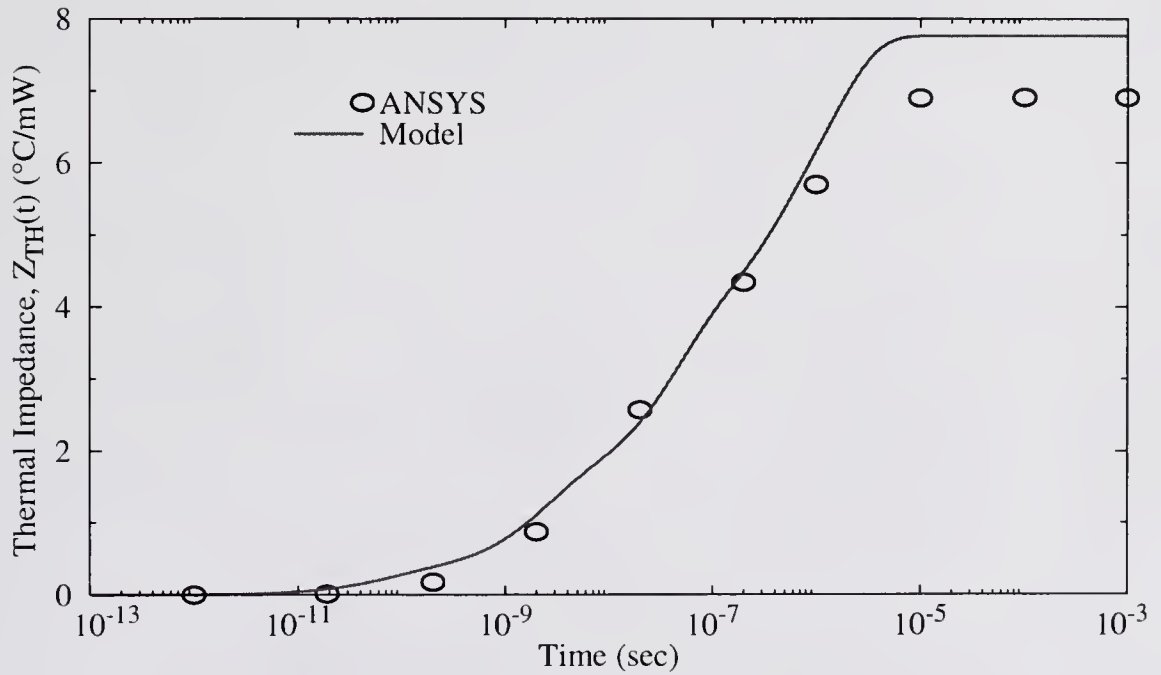
(b)

Figure 6.11　The transient thermal impedance simulated with ANSYS and calculated with the SOI MOSFET thermal impedance model for $W = 14\ \mu m$,　$L_d = 0.6\ \mu m$,　$d_d = 0.1\ \mu m$,　$d_{box} = 0.4\ \mu m\ W_m = 2\ \mu m$ and $d_m = 0.9\ \mu m$: a) For $L = 0.6\ \mu m$; b) for $L = 0.2\ \mu m$.

attributed to the lack of modeling of the contact structure itself. The model effectively assumes that the thermal resistance and capacitance of the contact structure is negligible. For devices with a large contact area, this assumption is valid. However, for devices with a relatively small contact area, the effectiveness of the drain/source interconnects as heat conductance paths is diminished. The temperature in such devices will rise quicker and to a greater magnitude than predicted by the model. The extent of the discrepancy can be reduced by tuning the width (and/or thickness) of the drain/source interconnects to reduce their effect on the thermal impedance.

Equation (6.47), in the complete form, was also compared to measured thermal impedance data extracted by Lee and Fox using transient drain current measurements [Lee95]. Figure 6.12 compares the measured and simulated data for a transient thermal impedance of an SOI MOSFET fabricated at Texas Instruments. Again, the model accurately predicts both the steady-state and transient temperature response. Part of the discrepancy at the earlier times is due to the finite rise time of the pulse used in the measurements. The model, as calculated for Figure 6.12, represents the temperature response to an ideal step in power dissipation, and will show a faster rise than the measurements.

6.6 Summary

Due to the low thermal conductivity of the silicon dioxide used for the buried oxide isolation, the thermal impedance of SOI MOSFET's is typically larger
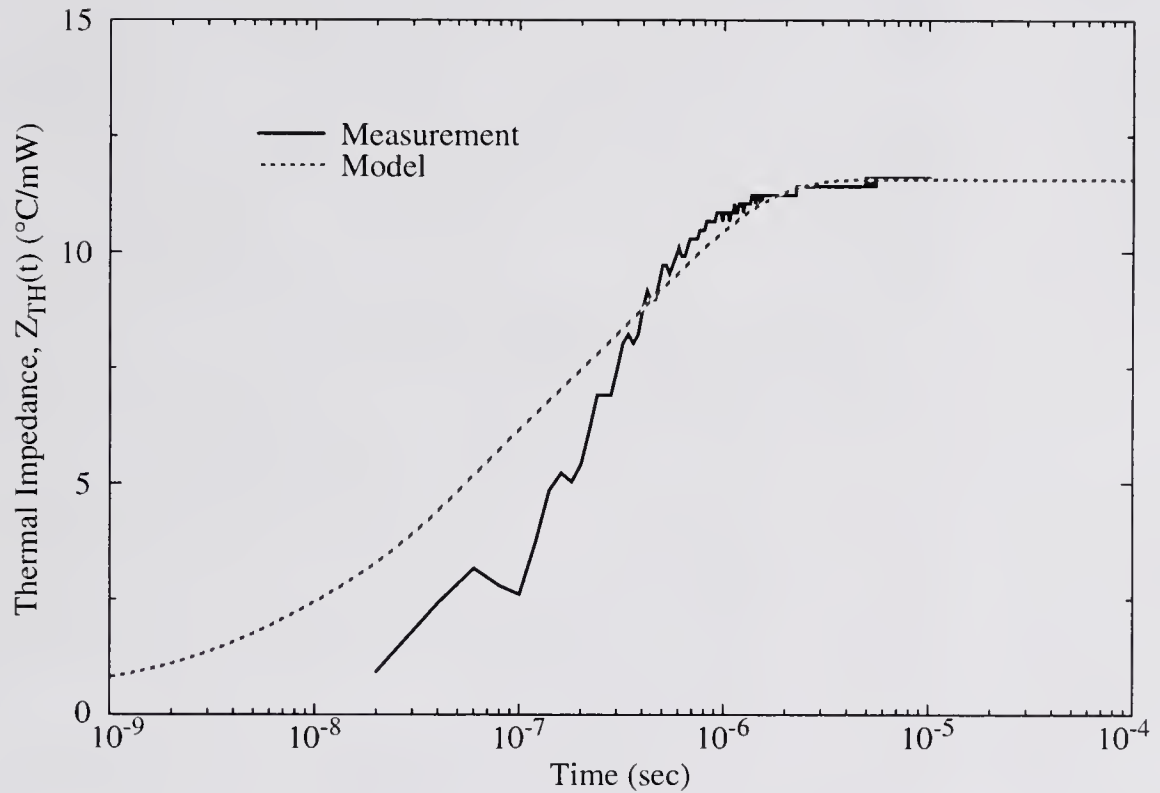
Figure 6.12   Measured and simulated data for the transient thermal impedance of a Texas Instruments SOI MOSFET with $W = 2.4\ \mu m$, $L = 1.8\ \mu m$, $d_d = 0.17\ \mu m$ and $d_{box} = 0.33\ \mu m$.

than that of their bulk counterparts. Therefore, self-heating effects can be enhanced and it is important to have a physical model that can predict the dynamic temperature rise. This chapter presented a physical thermal impedance model for SOI MOSFET's. The model was shown to agree reasonably well with both three-dimensional finite-element simulations and measurements.

CHAPTER 7
THE THERMAL IMPEDANCE PRE-PROCESSOR: TIPP

7.1 Introduction

The most common way to incorporate dynamic thermal effects into circuit simulators, using a thermal impedance to model the effective temperature rise in a transistor, is described in Chapter One. The analogy between electrical and thermal behavior can be used to represent the thermal impedance with a thermal equivalent circuit consisting of discrete resistances and capacitances. Consequently, thermal equivalent circuits provide a suitable means for efficient DC, AC and transient electrothermal circuit simulation.

For steady-state heat conduction, there is a simple correspondence between the thermal impedance and the thermal equivalent circuit: the sum of the resistance components in the thermal equivalent circuit is equal to the thermal spreading resistance. However, for a full dynamic thermal response, the correlation between the thermal impedance data and the component values of the thermal equivalent circuit is more complex. Thus, there is a need for a consistent and efficient approach for generating the component values from dynamic thermal impedance data.

This chapter presents a computer program that implements a systematic approach for calculating the component values of the thermal equivalent circuit. The

general operation of the program is discussed briefly. The numerical algorithms used to generate the thermal equivalent component values from thermal impedance data are examined in some detail. Examples of calculated thermal equivalent circuits are presented and a method for interfacing the program with electrothermal circuit simulators (ETCS's) is discussed.

7.2 A Description of TIPP

The Thermal Impedance Pre-Processor (TIPP) was developed to facilitate the use of ETCS's for modeling self-heating effects in semiconductor circuits. TIPP is a stand-alone software package for the specific purpose of providing an ETCS with access to thermal impedance models and the component values for their thermal equivalent circuits. The evolution of TIPP as a pre-circuit-simulation processor serves three purposes. First, the complexities of modifying an ETCS program to incorporate a thermal impedance model are avoided. The modifications would be specific to each circuit simulator and each thermal impedance model; therefore, the modifications would have to be repeated for each implementation. Second, the numerical burden of calculating the component values for the thermal equivalent circuit is removed from the circuit simulator. Third, TIPP was designed to provide an open frame-work for the generation of physics-based thermal impedance models and their thermal equivalent circuits; and therefore, gives a modifiable platform which can be accessed by any ETCS. TIPP is written in a modular format using the C programming language. The program structure allows for the easy addition of new

thermal impedance models or new numerical algorithms. At the present time, TIPP contains the physics-based models that are derived in the preceding chapters. While physics-based models offer the predictive capabilities that are useful in a development environment, they can not always provide the necessary accuracy required by a design environment. In such cases, the precision of optimized empirical data is essential. The numerical methods that TIPP uses to generate the thermal equivalent circuit are independent of the data source; therefore, TIPP can use either measured thermal impedance data or the predictive physical models to generate the component values for the thermal equivalent circuit.

The output from TIPP is in the form of a netlist which contains the components for any necessary thermal equivalent circuit(s). The generated thermal equivalent netlist can then be incorporated into an existing circuit file in preparation for an electrothermal simulation. The TIPP program structure is simple and was designed to provide results without drastically increasing the complexity of circuit simulation. TIPP is invoked from the command line using the parameters in Table 7.1.

Table 7.1 TIPP command-line parameters

Name	Description
filename	Name of file that contains input for TIPP
−sp	Option flag to instruct TIPP to use SPICE format for output file (filename.sp)
−sa	Option flag to instruct TIPP to use SABER format for output file(s) (filename.sin)

The basic operation of TIPP is outlined by the flowchart shown in Figure 7.1. The input for TIPP is contained in a text file that is composed of a set of instruction parameters, which are defined in Table 7.2.

Table 7.2 TIPP input parameters

Name	Description	Value
DEV	Specifies which thermal impedance model to calculate	BJT, SOIBJT, MOS, SOIMOS
DATA	Specifies whether to use a physics-based model or measured data to generate the thermal impedance data, where mfilename is name of the file that contains the measured data	PHYS or MEAS mfilename
NP	Specifies the number of poles for thermal equivalent circuit	1-5

If a physical thermal impedance model is used to generate the thermal equivalent circuit, the input file will also contain a list of geometric parameters that describe the layout of the device. Each model has a specific set of parameters that correspond to the necessary model variables. If measured data is used to generate the thermal equivalent circuit, the input file must contain the path and the name of the file that contains the data. An example of a typical TIPP input file is shown in Figure 7.2. After TIPP parses the input file, the thermal impedance data is either generated or imported. TIPP then passes the data to the algorithms that calculate the component values for a thermal equivalent circuit.
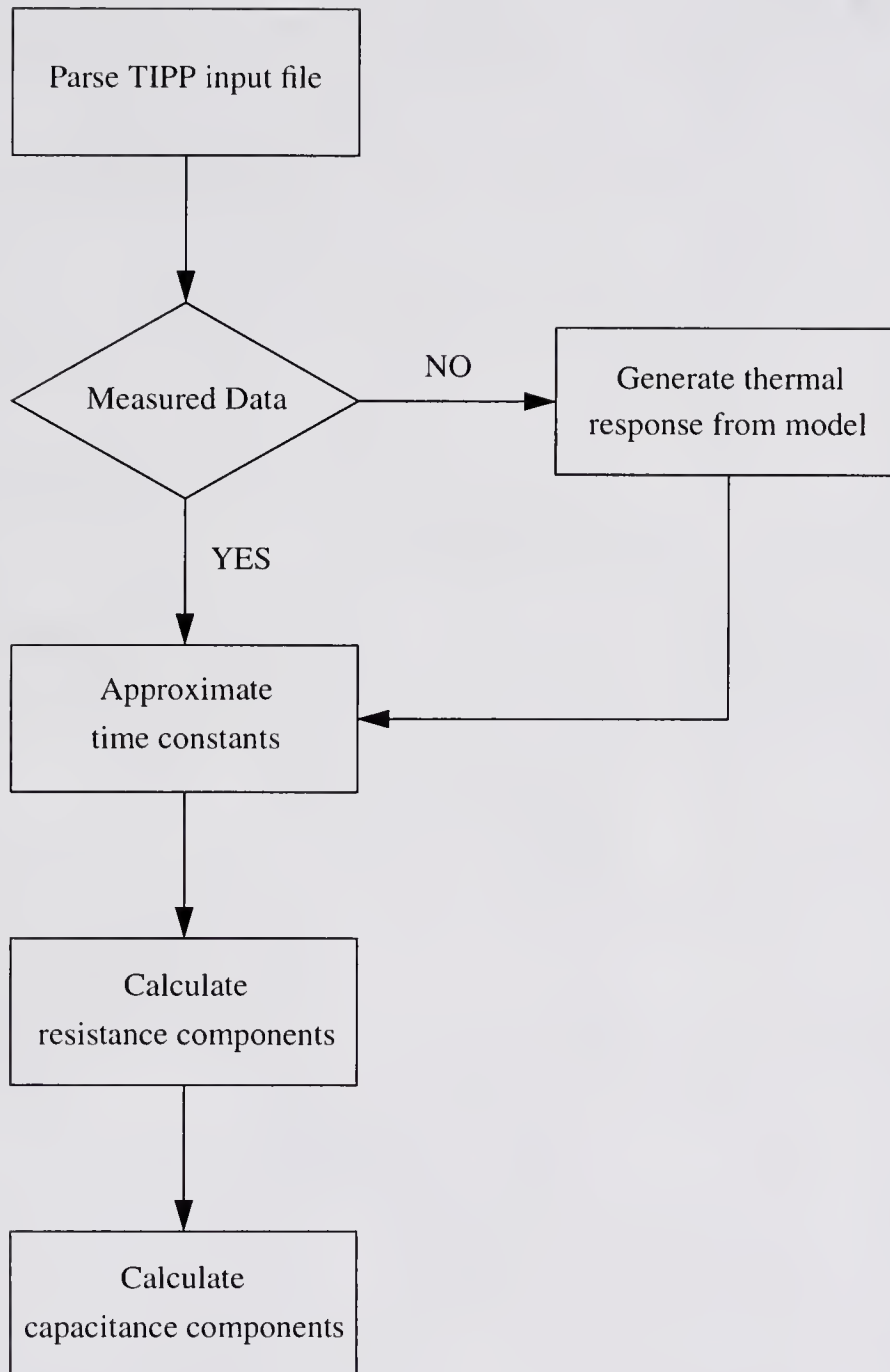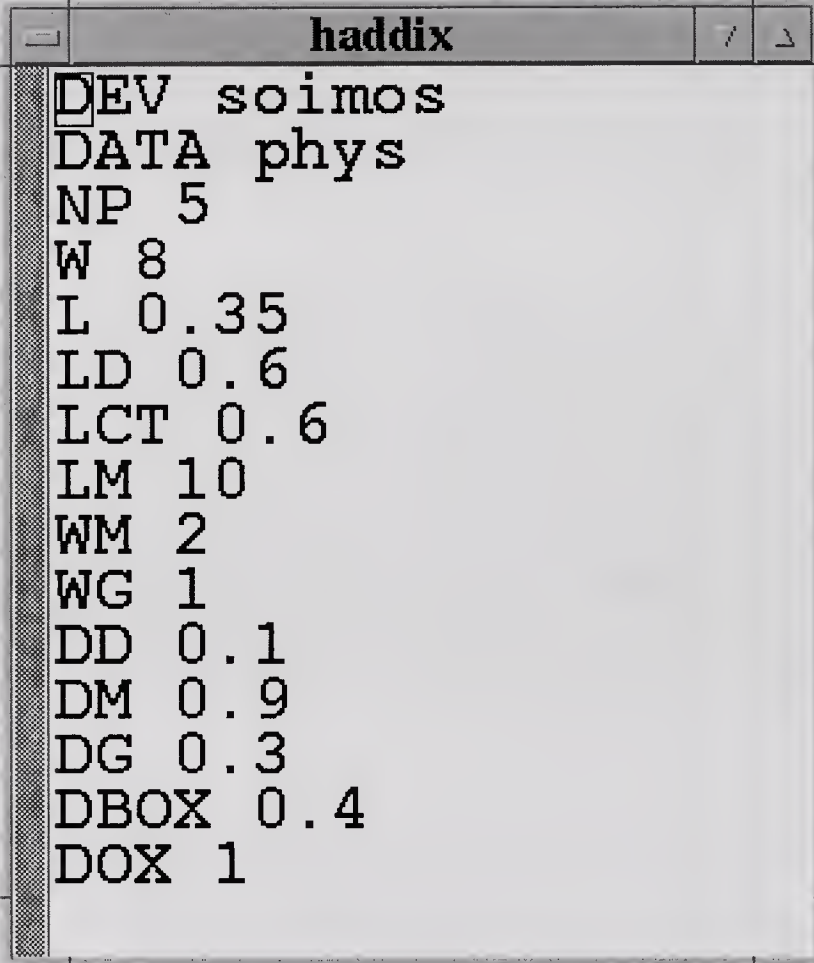
Figure 7.1   A flowchart illustrating the general operation of the TIPP software package

```
haddix

DEV soimos
DATA phys
NP 5
W 8
L 0.35
LD 0.6
LCT 0.6
LM 10
WM 2
WG 1
DD 0.1
DM 0.9
DG 0.3
DBOX 0.4
DOX 1
```

Figure 7.2   Example of a TIPP input file for an SOI MOSFET. The first three lines contain the TIPP-specific parameters for the simulation. The other parameters are used to describe the geometry of the transistor.

7.3 Generation of Thermal Equivalent Circuits

In Chapter One, the representation of a thermal impedance using circuit elements was discussed, and for this work, a Foster network was chosen to model the normalized temperature response. The components for the thermal equivalent circuit are generated by numerically fitting the multiple-pole network response to thermal impedance data; which TIPP assumes is in the form of a normalized (units of °C/W) transient step response or a normalized frequency response. Therefore, the network response of the thermal equivalent circuit must be derived in both the frequency and time domain.

The voltage generated across the thermal equivalent circuit can be expressed as

$$V(s) = I(s) \cdot Z_{TH}(s), \tag{7.1}$$

where $V(s)$ represents the effective temperature rise of the device and $I(s)$ represents the power dissipation. The normalized temperature response for an impulse in the power, where the Laplace transform of the impulse function is $\pounds[\delta(t)] = 1$, is simply equivalent to the impedance of the thermal equivalent circuit. Thus, for an n-stage Foster network, the thermal impedance can be expressed as

$$Z_{TH}(s) = \frac{r_{th1}}{1 + s/\omega_1} + \frac{r_{th2}}{1 + s/\omega_2} + \cdots + \frac{r_{thn}}{1 + s/\omega_n}, \tag{7.2}$$

where r_{thj} is the j-th thermal resistance component, $\sum r_{thj} \cong R_{TH}$ is the thermal spreading resistance and ω_j is the j-th pole. Equation (7.1) yields the temperature

response to a unit step function, $U(t)$, when $I(s) = 1/s$. Therefore, the time-domain thermal step response can be derived by multiplying (7.2) by $1/s$ and taking the inverse Laplace transform. The resulting expression is

$$Z_{TH}(t) = r_{th1}\left[1 - \exp\left(\frac{-t}{\tau_1}\right)\right] + r_{th2}\left[1 - \exp\left(\frac{-t}{\tau_2}\right)\right] + \cdots + r_{thn}\left[1 - \exp\left(\frac{-t}{\tau_n}\right)\right] , \quad (7.3)$$

where τ_j is the j-th time constant and $\omega_j = 1/\tau_j$.

7.3.1 Approximation of the Thermal Equivalent Poles/Time Constants

The first step in generating the thermal equivalent circuit is to calculate the pole or time constant associated with each stage of the thermal equivalent circuit. The initial approach used to calculate each pole/time constant was nonlinear optimization. The efficiency of the optimization algorithm was strongly dependent on the number of poles/time constants and the accuracy of the initial guess for the poles/time constants. However, the accuracy of the thermal equivalent circuit fit was found to be relatively insensitive to the exact location of the poles/time constants. Figure 7.3 shows an example of the variation in the fit accuracy that can occur when the locations of the poles/time constants are shifted. The variation in the fit accuracy was typically less than five percent for a ten percent change in any pole/time constant, and was never greater than ten percent. The results are typical for each thermal impedance model over most practical device geometries. Consequently, the insensitivity of the fit accuracy to the pole/time constant location can be exploited to develop a heuristic approximation for calculating the location.
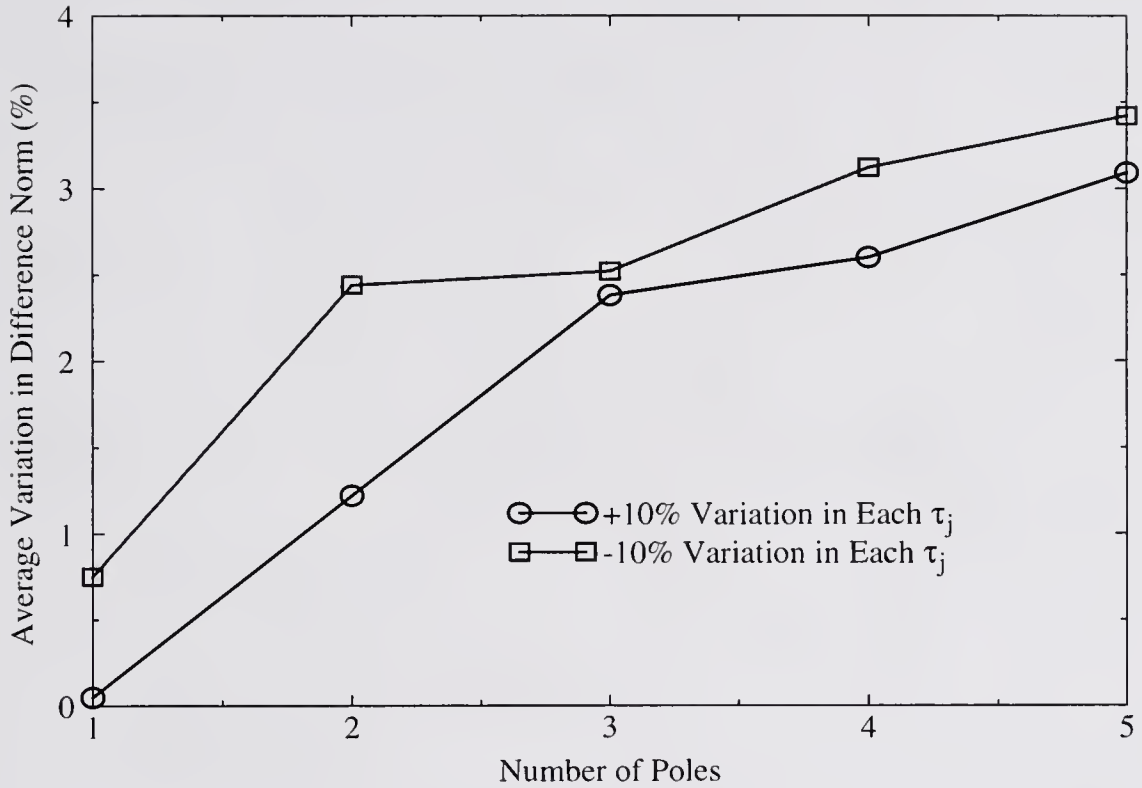
Figure 7.3   Example of the variation in the accuracy of a thermal equivalent circuit due to changes in the location each time constant. The bulk bipolar model is used to generate the thermal impedance data. The time constants of the thermal equivalent circuit are independently varied by ten percent in either direction. The difference norm between the thermal impedance data (TID) and the thermal equivalent circuit (TEC), $\sqrt{(TID_1 - TEC_1)^2 + (TID_2 - TEC_2)^2 + \cdots + (TID_m - TEC_m)^2}$, is calculated for m discrete data points. The variation in the difference norm is averaged over each time constant.

The algorithm for approximating the poles/time constants employs a binary search procedure that "scans" the thermal impedance magnitude for specific points. Figure 7.4 shows a graphical illustration of the algorithm when used to approximate five poles of a thermal equivalent circuit from a frequency-domain thermal impedance. For each stage of the thermal equivalent circuit, a single point is found which corresponds to a certain percentage, $0 < p_j < 1$, of the impedance maximum, max. The corresponding frequencies/times at which these points occur are used to approximate the poles/time constants. The value used for each percentage, p_j, was determined from the results of the original optimization algorithm performed on the bulk bipolar thermal impedance model. The model was evaluated for a wide range of device geometries and the final value for each percentage was taken as the average over this range. Though the values for the percentages were calculated for a specific model, they tend to be independent of the thermal impedance data source and provide accurate fits in most cases.

The aforementioned approach for determining the location of the poles/ time constants for a thermal equivalent circuit is simple and more efficient than numerical optimization. However, this technique is not completely robust, and could produce erroneous results if there is sufficient noise is present in the thermal impedance data. Notwithstanding, the open program structure of TIPP would allow the use of possibly more robust techniques such as moment matching [Pil94].
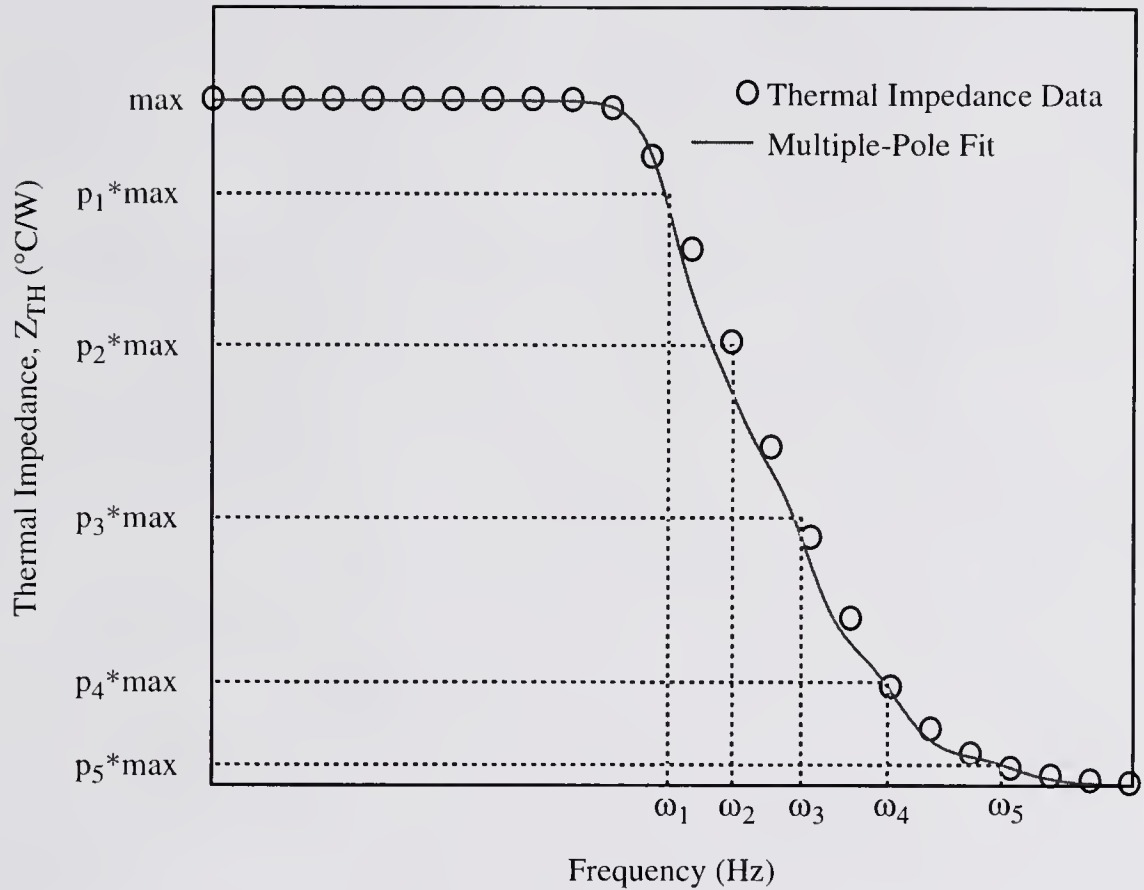
Figure 7.4   Illustration of the algorithm that TIPP uses to approximate the poles/ time constants for a thermal equivalent circuit. In this example, a five-pole thermal equivalent circuit is fit to a frequency-domain thermal impedance. The magnitude of the thermal impedance is scanned for specific points that correspond to certain percentages, p_j, of the impedance maximum, max. The frequencies at which these points occur are used to approximate the poles, ω_j, of the thermal equivalent circuit.

7.3.2 Calculation of the Thermal Equivalent Components

After the necessary poles/time constants are predicted, the resistance and capacitance components of the thermal equivalent circuit can be calculated. Equations (7.2) and (7.3) show that the response of the thermal equivalent circuit is linearly proportional to the resistance components. Thus, the resistance components can be efficiently derived by minimizing the discrete least-squares error between the thermal impedance data and the response of the Foster network.

A constraint is placed upon the thermal resistance components such that for an n-stage thermal equivalent circuit

$$r_n = \text{ssresp} - (r_1 + r_2 + \cdots + r_{(n-1)}) , \qquad (7.4)$$

where ssresp is the steady-state value of the thermal impedance data; therefore, the system of unknowns is reduced from n to $(n-1)$ independent variables. The least-squares approach along with (7.4) yields the following system of linear equations

$$
\begin{bmatrix}
\Lambda_{11} & \Lambda_{12} & \cdots & \Lambda_{1(n-1)} \\
\Lambda_{21} & \Lambda_{22} & \cdots & \Lambda_{2(n-1)} \\
\cdot & \cdot & & \cdot \\
\cdot & \cdot & & \cdot \\
\cdot & \cdot & & \cdot \\
\Lambda_{(n-1)1} & \Lambda_{(n-1)2} & \cdots & \Lambda_{(n-1)(n-1)}
\end{bmatrix}
\cdot
\begin{bmatrix}
r_{th1} \\
r_{th2} \\
\cdot \\
\cdot \\
\cdot \\
r_{th(n-1)}
\end{bmatrix}
=
\begin{bmatrix}
\Gamma_1 \\
\Gamma_2 \\
\cdot \\
\cdot \\
\cdot \\
\Gamma_{(n-1)}
\end{bmatrix}
\qquad (7.5)
$$

where

$$\Lambda_{jk} = \sum_{i=1}^{m} \left[\exp\left(\frac{-t_i}{\tau_n}\right) - \exp\left(\frac{-t_i}{\tau_j}\right) \right] \cdot \left[\exp\left(\frac{-t_i}{\tau_n}\right) - \exp\left(\frac{-t_i}{\tau_k}\right) \right] \tag{7.6}$$

and

$$\Gamma_j = \sum_{i=1}^{m} \left[\exp\left(\frac{-t_i}{\tau_n}\right) - \exp\left(\frac{-t_i}{\tau_j}\right) \right] \cdot \left\{ resp_i - ssresp\left[1 - \exp\left(\frac{-t_i}{\tau_n}\right) \right] \right\} \tag{7.7}$$

when using (7.3) to fit a normalized thermal step response, resp, with m discrete data points. If resp corresponds to a normalized thermal frequency response, equation (7.2) is used to fit the data so that

$$\Lambda_{jk} = \sum_{i=1}^{m} \left\{ \left| \frac{1}{1 + s_i/\omega_j} \right| - \left| \frac{1}{1 + s_i/\omega_n} \right| \right\} \cdot \left\{ \left| \frac{1}{1 + s_i/\omega_k} \right| - \left| \frac{1}{1 + s_i/\omega_n} \right| \right\} \tag{7.8}$$

and

$$\Gamma_j = \sum_{i=1}^{m} \left\{ \left| \frac{1}{1 + s_i/\omega_j} \right| - \left| \frac{1}{1 + s_i/\omega_n} \right| \right\} \cdot \left\{ resp_i - ssresp \cdot \left| \frac{1}{1 + s_i/\omega_n} \right| \right\} . \tag{7.9}$$

The n resistance components can be calculated by solving (7.5) and (7.4). Since the resulting coefficient matrix is never larger than 4×4, Gaussian elimination and back-substitution are used to solve the system of linear equations.

After the resistance components are determined, the capacitance components can be calculated using the following equation

$$c_{thj} = \tau_j/r_{thj} = 1/(\omega_j r_{thj}).\tag{7.10}$$

Examples of complete thermal equivalent circuit fits are shown in Figure 7.5. The thermal equivalent circuits show good agreement with the thermal impedance data generated from both measurement and a physics-based model. The plots also illustrate how the distributed nature of the thermal impedance is more accurately represented by thermal equivalent circuits using multiple poles.

7.4 The Interface Between TIPP and Circuit Simulators

The last operation of the TIPP program is to output the calculated resistance and capacitance components of a thermal equivalent circuit. Since TIPP is intended to function as a companion to any given ETCS, the format of the output should be consistent with the syntax of that simulator. Therefore, the interface that TIPP uses to communicate with an ETCS should be flexible and convenient.

The role of a pre-processor was chosen so that TIPP's own program structure would be independent of ETCS program syntax. While this approach is not as immediate as directly coupling TIPP to an ETCS, it avoids the unwieldy modifications, to both circuit simulator and TIPP, that would be required to implement a direct interface. Since it is simple to modify or extend the TIPP program alone, its output can be constructed in a format that is compatible with any ETCS. As shown in Table 7.1, TIPP's output can be formatted for either a SPICE-type or the SABER [Sab95] circuit simulator. This output is written to a text file, or group of text files, that can easily be incorporated into an existing circuit file. If a specific
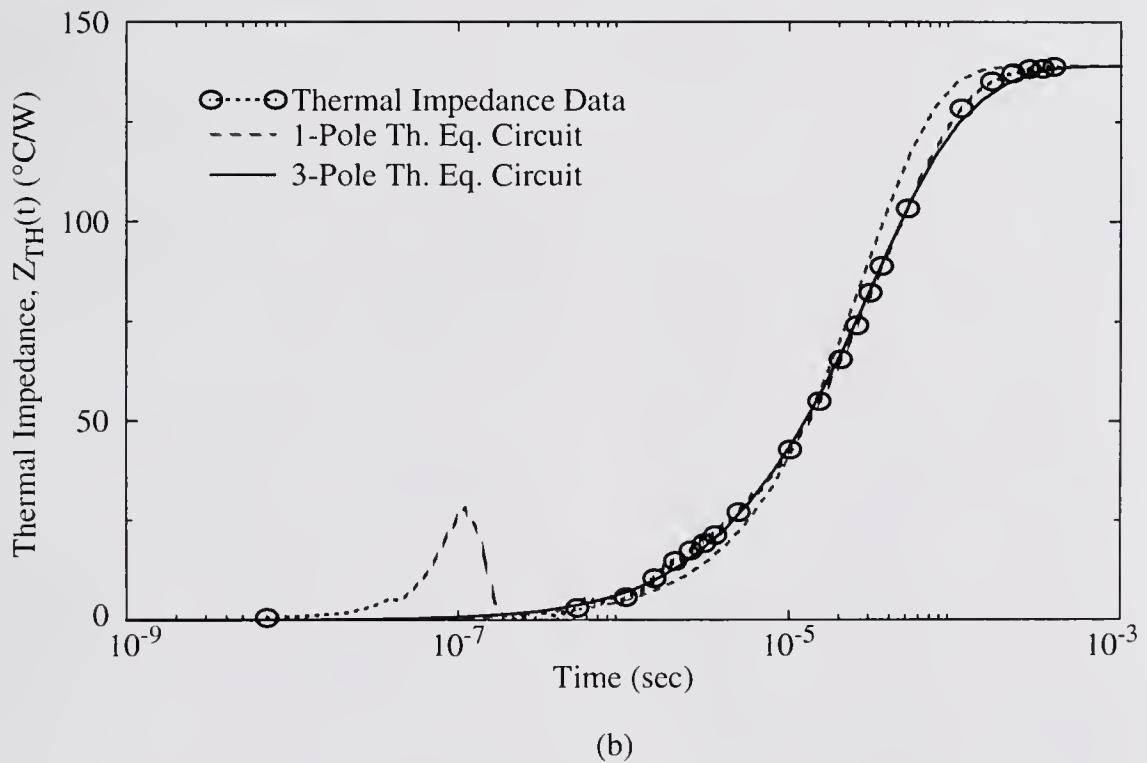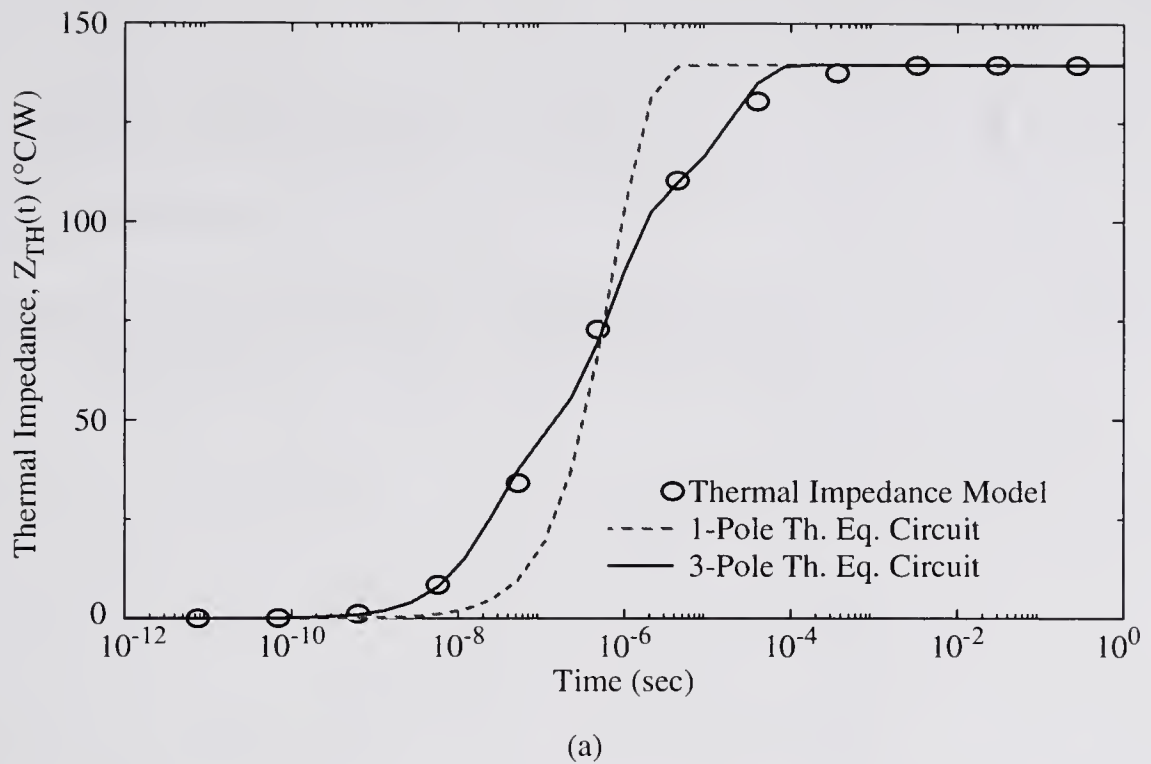
Figure 7.5   Plots of thermal equivalent circuit representations of thermal impedance data: a) Thermal impedance generated using a physical model; b) thermal impedance extracted from measurements.

output format is not selected, TIPP writes a simple list of the components for each thermal equivalent circuit.

For SPICE-type circuit simulators, TIPP writes a single output file (filename.sp) that contains a subcircuit netlist for each thermal equivalent circuit. For the MEBJT/MEHBT thermal impedance model, the output file will also contain a subcircuit for the complete device. The MEBJT/MEHBT subcircuit consists of the necessary transistor cards and elements to model each finger and each thermal-coupling network. For the SABER circuit simulator, TIPP writes one output file (filename.sin) for each generated thermal equivalent circuit. Each file contains the corresponding netlist in the proper SABER template format. If the MEBJT/MEHBT thermal impedance model is used to generate the thermal equivalent circuits, a separate output file is created for the MEBJT/MEHBT electrothermal model. As with the SPICE-formatted output, the MEBJT/MEHBT template contains the netlists for the necessary thermal-coupling networks. The formats used to generate the SPICE and SABER output files are shown in Figure 7.6 and Figure 7.7, respectively. If a circuit simulator does not use either of the mentioned formats, the TIPP program can be easily extended to provide the required output format.

7.5 Summary

In order to investigate thermal effects in circuits and devices, circuit simulators have been modified to account for dynamic temperature variations. The temperature response of a device can be modeled by its thermal impedance, which

```
* Subcircuit for Thermal Equivalent Circuit

.subckt 1  (np + 1) Z
rth1  1  2  <value>
cth1  1  2  <value>
       .

       .

       .

rth(np) np  (np + 1) <value>
cth(np) np  (np + 1) <value>
.ends Z




* Subcircuit for MEBJT/MEHBT

.subckt 1  2  3  4 MEBJT
<transistor elements for each emitter finger>

* Thermal Coupling Networks
xs1 os  (os + 1) ZS
ec12 (os + 1)  (os + 2)  ncp  ncn  1
       .

       .

       .

ec1nf (os + nf − 1)  (os + nf)  ncp  ncn  1
vs1 (os + nf)  0 0

fp1 0  (os + nf + 1) vs1 1
xc21 (os + nf + 1)  (os + nf + 2) ZC1
       .

       .

       .

xcnf1 (os + 2nf − 1)  0 ZC(nf-1)
<repeated for each emitter finger>
.ends MEBJT
```

Figure 7.6   Template for TIPP output in SPICE format, where os is an arbitrary off-set for the node numbering.

```
# Template for Thermal Equivalent Circuit

template Z 1 (np + 1)
{
r.rth1 1  2 = value
c.cth1 1  2 = value

       .

       .

       .

r.rth(np) np (np + 1) = value
c.cth(np) np (np + 1) = value
}


# Template for MEBJT

template MEBJT 1 2 3 4
{
<transistor elements for each emitter finger>

# Thermal Coupling Networks
ZS.zs1 os (os + 1)
vcvs.tc12 ncp  ncn (os + 1) (os + 2) = k = 1

       .

       .

       .

vcvs.tc1nf ncp  ncn (os + nf − 1) (os + nf) = k = 1
v.vs1 (os + nf) 0 = 0

cccs.p1 i(v.vs1) 0 (os + nf + 1) = k = 1
ZC1.zc21 (os + nf + 1) (os + nf + 2)

       .

       .

       .

ZC(nf-1).zcnf1 (os + 2nf − 1) 0
<repeated for each emitter finger>
}
```

Figure 7.7   Template for TIPP output in SABER format, where os is an arbitrary off-set for the node numbering.

can be derived from physical models or extracted from measurements. Multiple-pole thermal equivalent circuits provide an efficient method for incorporating thermal impedance data into a circuit simulation. The TIPP computer program was developed as a general frame-work for thermal impedance modeling. TIPP generates thermal equivalent circuits using its internal physics-based models or imported measured data. The flexibility of the TIPP program structure allows new thermal impedance models to be added easily and simple interfacing with electrothermal circuit simulators.

CHAPTER 8
CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

8.1 Conclusions

The primary goal of this dissertation was to develop physical models that could predict the thermal impedance of semiconductor devices and also remain efficient enough to be used for circuit simulation. Four thermal impedance models were presented for different device structures, all of which satisfy the above-mentioned conditions. Each model is derived from the physical heat conduction equation, and is therefore capable of accurately predicting the thermal impedance based solely on material properties and device geometry. In addition, the models are in the form of compact analytical expressions which produce results quickly, therefore offering a more efficient alternative to costly numerical thermal simulations.

The development of thermal impedance models for BJT's, on both bulk and SOI substrates, was detailed in Chapters Two and Four. The bulk BJT model, originally developed by Joy and Schlig [Joy70], was extended to account for finite substrate thickness and devices with multiple emitters. The DIBJT model was derived to depict the complete transient temperature response and was then simplified to produce an additional model for the steady-state thermal resistance. Both models were augmented to account for the thermal resistance and heat capacity

of the emitter interconnect metallization. In addition, three-dimensional finite-element simulations were used to show that the effects of advanced dielectric isolation on the thermal impedance of bulk BJT's can be significant for highly-scaled transistors.

In Chapter Three, a circuit model for thermal coupling between devices in a common substrate was derived. The circuit representations were developed to facilitate the implementation of the multiple-emitter BJT thermal impedance model in circuit simulation; however, they are also applicable to the general case of thermal coupling between arbitrary devices. The thermal coupling model was then used to develop an extraction methodology for an efficient, lumped electrothermal model for multiple-emitter BJT/HBT's.

Thermal impedance models for MOSFET's, on both bulk and SOI substrates, were described in Chapters Five and Six. The bulk MOSFET model improved on the work by Sharma and Ramanathan [Sha83] by accounting for cross-width temperature gradients and the linear variation of the electric field along the length of the channel. Three-dimensional finite-element simulations were used to investigate the effects of device interconnects and isolation on the thermal impedance of bulk devices. The SOI MOSFET model was based on a modified version of the steady-state thermal resistance model developed by Goodson and Flik [Goo92]. The modified steady-state model was extended for time-dependent heat conduction to produce the dynamic thermal impedance model.

In Chapter Seven, the Thermal Impedance Pre-Processor (TIPP) was introduced. The TIPP software program was developed to provide efficient and

systematic algorithms for generating thermal impedance models and the thermal equivalent circuits necessary for electrothermal circuit simulations.

Finally, a number of the accomplishments that were the result of this research, have been documented in various journal and conference publications. These papers are cited, as follows, in the reference list at the end of this dissertation [Bro93, Bro97, Fox93a, Zwe95a, Zwe95b, Zwe96].

8.2 Recommendations for Future Work

The thermal impedance models presented in this dissertation can provide accurate predictions when used within the limits of their derivations. However, the robustness of the models, over a wider range of device structure types and sizes, can be enhanced further. Two areas are presented for future research that could improve the accuracy of the models. In addition, a third research topic concerning thermal coupling in SOI MOSFET circuits is suggested.

8.2.1 The Temperature Dependence of the Thermal Conductivity

The heat conduction equation used to derive the thermal impedance models in this dissertation assumes that the thermal conductivity is constant, and therefore independent of temperature. For small to moderate temperature excursions, this assumption is valid; however, for a large temperature rise (50-100 °C) above the reference ambient, the thermal conductivity can vary by as much as 31% [Gao89]. Such large temperature excursions in transistors are common in high-power

applications and therefore, the accuracy of the thermal impedance models could be improved by accounting for the temperature dependence of the thermal conductivity.

The variation of the thermal conductivity with temperature makes the heat conduction equation non-linear, such that

$$\nabla \bullet k(T) \nabla T + g = \rho \cdot c_p \frac{\partial T}{\partial t},$$ (8.1)

where the temperature dependence of the specific heat is assumed to be relatively weak compared to that of the thermal conductivity. The Kirchoff transformation [Joy75] linearizes (8.1) by using the following variable transformation

$$\theta = T_0 + \frac{1}{k_0} \int_{T_0}^{T(t)} k(T) \cdot dT,$$ (8.2)

where k_0 is the thermal conductivity at the reference ambient temperature, T_0. The relation between the linear and non-linear temperature variables can be examined by multiplying both sides of (8.2) by k_0, taking the gradient and then the divergence, resulting in

$$k_0 \nabla^2 \theta = \nabla \bullet k(T) \nabla T = \rho \cdot c_p \frac{\partial T}{\partial t} - g.$$ (8.3)

Consequently, if the temperature dependence of the thermal conductivity is known and can be expressed as an integrable equation, (8.2) can be used to correct the temperature calculated with the linear heat conduction equation.

The correction to the estimated temperature, accounting for the temperature dependence of the thermal conductivity, can be implemented in an electrothermal circuit simulator by adjusting the voltage generated at the temperature node. Preliminary attempts at such a modification have shown that the approach is viable. However, the implementation must de handled carefully to ensure consistency between the thermal and electrical solutions, and to reduce possible convergence problems. It is suggested that further research be conducted to ascertain the optimum approach for incorporating the Kirchoff transformation into an electrothermal circuit simulator.

8.2.2 Models for Thermal Effects Due to Advanced Isolation

The thermal impedance models for bulk BJT's and MOSFET's described in Chapters Two and Five, neglect the effects of the oxide isolation structures (e.g. LOCOS or trench) used to electrically isolate devices in a common substrate. Due to their low thermal conductivity, these isolation structures can significantly increase the thermal impedance of highly-scaled devices. Consequently, there can be substantial errors between the results predicted by the models and actual observed thermal impedances. Therefore, the bulk models should be modified to improve their accuracy for highly-scaled devices that are fabricated with advanced isolation structures.

By implementing the Green's function technique that is detailed in Chapter Four, both the bulk BJT and bulk MOSFET thermal impedance models can be modified to account for isolation structures. Adapting this analysis to the bulk

models would require that the physical device be divided into two domains, which can be thought of as intrinsic and extrinsic thermal regions. The intrinsic thermal domain can be approximated by a rectangular volume with boundaries defined at the surface of the substrate, at the depth of the isolation structure and at the interfaces between the intrinsic electrical device and the "side-walls" of the isolation structure. The extrinsic thermal domain would represent the exterior portions of the substrate that surround the isolation structure. The three-dimensional heat conduction equation would be solved within the intrinsic thermal domain. The thermal boundary conditions would be derived from the nature of the thermal conduction in the extrinsic thermal domain. The difficulty in this implementation, lies in the derivation of the boundary conditions. It is suggested that the ANSYS finite-element solver be employed to extensively simulate isolated device structures. From these simulations, reasonable assumptions can be formulated to simplify the heat conduction analysis in the exterior thermal domain for different isolation structures. Once the thermal analysis is simplified, boundary conditions for the intrinsic thermal domain can be deduced.

8.2.3 A Model for Thermal Coupling in SOI MOSFET Circuits

For MOSFET's fabricated in a common substrate using a bulk technology, thermal coupling between devices is due primarily to heat transport through the substrate. In such cases, the thermal impedances that describe the thermal interactions between devices can be modeled using the analysis similar to that in Chapter Two for MEBJT's. Due to the insulating properties of the buried and field

oxides, MOSFET's fabricated on an SOI substrate might be considered thermally isolated from one another. However, recent work by Tenbroek et al. [Ten96] showed that there can be thermal coupling between MOSFET's on a common SOI substrate. Since SOI devices are effectively isolated within the substrate, the dominant mechanism of thermal coupling between MOSFET's is the heat conduction through the device interconnects.

The effects of thermal coupling in an SOI MOSFET current mirror are illustrated in Figure 8.1; the circuit shown in the inset is modeled after the current mirrors measured by Tenbroek et al. [Ten96] and was simulated in SOISPICE [Fos95]. The reference device and device M1 are assumed to be located in close proximity of one another such that there is significant thermal coupling. Device M2 is assumed to be located a relatively large distance away from the other devices, and is therefore thermally isolated. The drain voltage of M2 is fixed while the drain voltage of M1 is allowed to vary. The drain current of device M2 should be equivalent to, or mirror, the drain current of the reference device and should be independent of V_{D1}. However, as the simulation shows, the drain current of M2 does depend on the drain voltage V_{D1}. This phenomenon can be explained by the thermal coupling between M1 and the reference device. As the power in M1 increases, the temperature in the reference device will increase. Since the drain current of the reference device is fixed, the gate voltage will have to increase to counter the effects of the reduced carrier mobility caused by the increase in temperature. Therefore, the gate voltage of M2 will also increase, resulting in an increase in drain current.
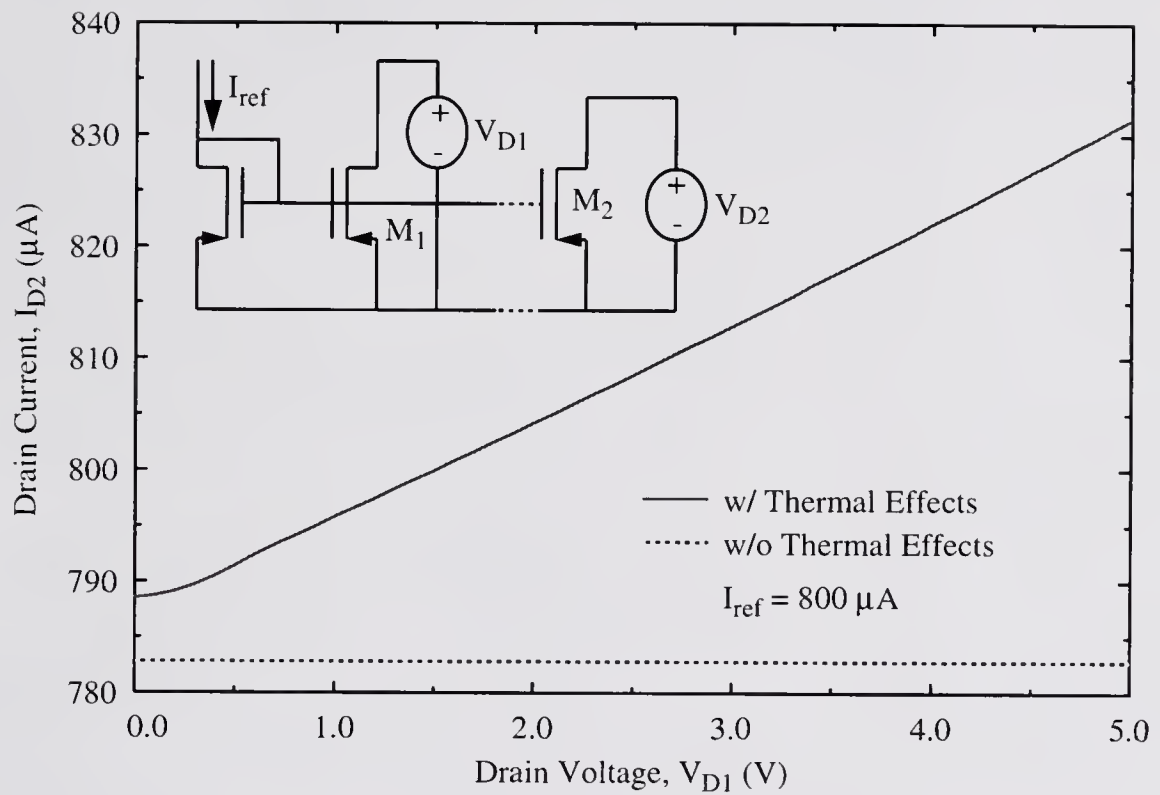
Figure 8.1   Simulation of an SOI MOSFET current mirror showing the effects of thermal coupling between devices.

A preliminary model for thermal-coupling impedance between SOI MOSFET's, which was used for the simulation in Figure 8.1, has been developed. However, the model is not complete. Two- and three-dimensional ANSYS simulations, used in conjunction with measurements of the test structures shown by Tenbroek et al. [Ten96], are suggested to further the development and validate the accuracy of the thermal-coupling impedance model. Once the model is complete, it can be used with a physics-based circuit simulator, such as SOISPICE [Fos95], to investigate the effect of thermal coupling in a wide range of SOI circuits.

REFERENCES

[Ans96]   ANSYS User's Manual, Revision 5.3, SAS IP, Inc., Houston, PA, June 1996.

[Bau93]   Peter Baureis and Dieter Seitzer, "Parameter Extraction for HBT's Temperature Dependent Large Signal Equivalent Circuit Model," in Proceedings GaAs IC Symposium, San Jose, CA, pp. 263-266, 1993.

[Bau94]   Peter Baureis, "Electrothermal Modeling of Multi-Emitter Heterojunction Bipolar Transistors (HBTs)," in Proceedings International Workshop on Integrated Nonlinear Microwave and Millimeter Circuits, Duisburg, Germany, pp. 145-148, 1994.

[Ber91]   Michael Berger and Zhiqin Chai, "Estimation of Heat Transfer in SOI-MOSFET's," IEEE Transactions on Electron Devices, vol. 38, no. 4, pp. 871-875, April 1991.

[Bie95]   Juergen Bielefeld, Geor Pelz, Hans Bernd Abel and Gunter Zimmer, "Dynamic SPICE-Simulation of the Electrothermal Behavior of SOI MOSFET's," IEEE Transactions on Electron Devices, vol. 42, no. 11, pp. 1968-1974, November 1995.

[Bro93]   Jonathan S. Brodsky, David T. Zweidinger and Robert M. Fox, "Physics-Based Multiple-Pole Models for BJT Self-Heating," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 249 -252, 1993.

[Bro97]   Jonathan S. Brodsky, Robert M. Fox, David T. Zweidinger and Surya Veeraraghavan, "A Physics-Based, Dynamic Thermal Impedance Model for SOI MOSFET's," IEEE Transactions on Electron Devices, vol. 44, no. 6, pp. 957-964, June 1997.

[Cav93]   Anthony L. Caviglia and Agis A. Iliadis, "Linear Dynamic Self-Heating in SOI MOSFET's," IEEE Electron Device Letters, vol. 14, no. 3, pp. 133-135, March 1993.

[Cav95]   Anthony Caviglia and Agis A. Iliadis, "A Large-Signal Model for SOI MOSFET's Including Dynamic Self-Heating Effects," in Proceedings IEEE International SOI Conference, Tucson, AZ, pp. 16-17, October 1995.

[Che95]   Yu-Guang Chen, Shyh-Yih Ma, James B. Kuo, Zhiping Yu and Robert W. Dutton, "An Analytical Drain Current Model Considering Both electron and Lattice Temperatures Simultaneously for Deep Submicron Ultrathin SOI NMOS Devices with Self-Heating," IEEE Transactions on Electron Devices, vol. 42, no. 5, pp. 899-906, May 1995.

[Dav92]   C. Davis, G. Bajor, J. Butler, T. Crandell, J. Delgado, T. Jung, Y. Khajeh-Noori, B. Lomenick, V. Milam, H. Nicolay, S. Richmond and T. Rivoli, "UHF-1: A High Speed Complementary Bipolar Analog Process on SOI," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 260-263, 1992.

[Daw94]   Dale E. Dawson, "Thermal Modeling, Measurements and Design Considerations of GaAs Microwave Devices," in Proceedings GaAs IC Symposium, Philadelphia, PA, pp. 285-290, 1994.

[Del91]   V. dela Torre, J. Foerstner, B. Lojek, K. Sakamoto, S. L. Sundaram, N. Tracht, B. Vasquez and P. Zdebel, "MOSAIC V - A Very High Performance Bipolar Technology," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 21-24, 1991.

[Fei92]   S. Feindt, J-J. J. Hajjar, J. Lapham and D. Buss, "XFCB: A High Speed Complementary Bipolar Process on Bonded SOI," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 264-267, 1992.

[Fos95]   J. G. Fossum, R. M. Fox and D. T. Zweidinger, "SOI-SPICE-4 (Version 4.1) Silicon-On-Insulator MOSFET SPICE with Self-Heating," Programmer's Reference Manual, University of Florida Department of Electrical and Computer Engineering, 1995.

[Fox91a]   Robert M. Fox and Sang-Gug Lee, "Predictive Modeling of Thermal Effects in BJTs," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 89-92, 1991.

[Fox91b]   Robert M. Fox and Sang-Gug Lee, "Scalable Small-Signal Model for BJT Self-Heating," IEEE Electron Device Letters, vol. 12, no. 12, pp. 649-651, December 1991.

[Fox93a]   R. M. Fox and J. S. Brodsky, "Effects of Self-Heating-Induced Negative Output Conductance in SOI Circuits," in Proceedings IEEE International SOI Conference, Palm Springs, CA, pp. 152-153, 1993.

[Fox93b]   Robert M. Fox, Sang-Gug Lee and David T. Zweidinger, "The Effects of BJT Self-Heating on Circuit Behavior," IEEE Journal of Solid-State Circuits, vol. 28, no. 6, pp. 678-685, June 1993.

[Fuk76]   Kiyoshi Fukahori and Paul R. Gray, "Computer Simulation of Integrated Circuits in the Presence of Electrothermal Interaction," IEEE Journal of Solid-State Circuits, vol. SC-11, no. 6, pp. 834-846, December 1976.

[Gan92]   P. R. Ganci, J-J. J. Hajjar, T. Clark, P. Humphries, J. Lapham and D. Buss, "Self-Heating in High Performance Bipolar Transistors Fabricated on SOI Substrates," in Technical Digest IEEE International Electron Devices Meeting, San Francisco, CA, pp. 417-420, 1992.

[Gao89]   Guang-Bo Gao, Ming-Zhu Wang, Xiang Gui and Hadis Morkoc, "Thermal Design Studies of High-Power Heterojunction Bipolar Transistors," IEEE Transactions on Electron Devices, vol. 36, no. 5, pp. 854-863, May 1989.

[Goo92]   K. E. Goodson and M. I. Flik, "Effect of Microscale Thermal Conduction on the Packing Limit of Silicon-On-Insulator Electronic Devices," IEEE Transactions on Components, Hybrids, and Manufacturing Technology, vol. 15, no. 5, pp.715-722, October 1992.

[Goo95]   K. E. Goodson, M. I. Flik, L. T. Su and D. A. Antoniadis, "Prediction and Measurement of Temperature Fields in Silicon-On-Insulator Electronic Circuits," ASME Journal of Heat Transfer, vol. 117, pp. 574-581, August 1995.

[Gra93]   Paul R. Gray and Robert G. Meyer, <u>Analysis and Design of Analog Integrated Circuits</u>, Third Edition, John Wiley & Sons, Inc., New York, 1993.

[Gre93]   Keith R. Green, <u>A Model of the Short-Channel Metal-Oxide-Semiconductor Field-Effect Transistor for Pragmatic Mixed-Mode Device/Circuit Simulation</u>, Ph. D. Dissertation, University of Florida, August 1993.

[Hat95]   Ryo Hattori, Teruyuki Shimura, Manabu Kato, Takuji Sonoda and Saburo Takamiya, "Three-Dimensional Modeling of Thermal Flow in Multi-Finger High Power HBTs," in Proceedings IEEE MTT-S International Microwave Symposium, Orlando, FL, vol. 2, pp. 461-464, 1995.

[Hir93]   I. Hirsch, E. Berman and N. Haik, "Thermal Resistance Evaluation in 3-D Thermal Simulation of MOSFET Transistors," Solid-State Electronics, vol. 36, no. 1, pp. 106-108, January 1993.

[Hsp92]   HSPICE User's Manual, Version H92, Meta-Software, Inc.,Campbell, CA, 1992.

[Jen95]   K. A. Jenkins and J. Y. -C. Sun, "Measurement of I-V Curves of Silicon-on-Insulator (SOI) MOSFET's Without Self-Heating," IEEE Electron Device Letters, vol. 16, no. 4, pp. 145-147, April 1995.

[Jeo89]   Hanggeun Jeong and Jerry G. Fossum, 'A Charge-Based Large-Signal Bipolar Transistor Model for Device and CIrcuit Simulation," IEEE Transactions on Electron Devices, vol. 36, no. 1, pp. 124-131, January 1989.

[Jer93]   R. C. Jerome, I. R. C. Post, P. G. Travnicek, G. M. Wodek, K. E. Huffstater and D. R. Williams, "ACUTE: A High Performance Analog Complementary Polysilicon Emitter Bipolar Technology Utilizing SOI/Trench Full Dielectric Isolation," in Proceedings IEEE International SOI Conference, Palm Springs, CA, pp. 100-101, 1993.

[Jom95a]   J. Jomaah, G. Ghibaudo and F. Balestra, "Analysis and Modeling of Self-Heating Effects in Thin-Film SOI MOSFETs as a Function of Temperature," Solid-State Electronics, vol. 38, no. 3, pp. 615-618, March 1995.

[Jom95b]   J. Jomaah, G. Ghibaudo, F. Balestra and J. L. Pelloie, "Impact of Self-Heating on the Design of SOI Devices Versus Temperature," in Proceedings IEEE International SOI Conference, Tucson, AZ, pp. 114-115, October 1995.

[Joy70]   Richard C. Joy and E. S. Schlig, "Thermal Properties of Very Fast Transistors, IEEE Transactions on Electron Devices, vol. ED-17, no. 8, pp. 586-594, August 1970.

[Joy75]   W. B. Joyce. "Thermal Resistance of Heat Sinks with Temperature-Dependent Conductivity," Solid-State Electronics, vol. 18, pp. 321-322, 1975.

[Kag94]   A Kager, J. J. Liou, L. L. Liou and C. Huang, "A Semi-Numerical Model for Multi-Emitter Finger AlGaAs/GaAs HBTs," Solid-State Electronics, vol. 37, no. 11, pp. 1825-1832, November 1994.

[Klo93]   H. Klose, R. Lachner, K. R. Schon, R. Mahnkopf, K. H. Malek, M. Kerber, H. Braun, A. v. Felde, J. Popp, O. Cohrs, E. Bertagnolli and P. Sehrig, "B6HF: A 0.8 Micron 25 GHz/25 ps Bipolar Technology for Mobile Radio and Ultra Fast Data Link IC-Products," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 125-127, 1993

[Koe94]   Eric Koenig, Jurgen Schneider, Ulrich Seiler and Uwe Erben, "Thermally Induced Current Constriction in III-V Heterojunction Bipolar Transistors," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 127-130, 1994.

[Lee93]   M. S. L. Lee, W. Redman-White, B. M. Tenbroek and M. Robinson, "Modeling of Thin Film SOI Devices for Circuit Simulation Including Per-Instance Dynamic Self-Heating Effects," in Proceedings IEEE International SOI Conference, Palm Springs, CA, pp. 150-151, 1993.

[Lee95]   T. -Y. Lee and R. M. Fox, "Extraction of Thermal Resistance for Fully-Depleted SOI MOSFET's," in Proceedings IEEE International SOI Conference, Tucson, AZ, pp. 78-79, October 1995.

[Lee96]   Tzung-Yin Lee, Model Enhancement and Parameter Extraction for the MMSPICE/QBBJT Model, Ph. D. Dissertation, University of Florida, December 1996.

[Lia94]   Minchang Liang and Mark E. Law, "Influence of Lattice Self-Heating and Hot-Carrier Transport on Device Performance," IEEE Transactions on Electron Devices, vol. 41, no. 12, pp. 2391-2398, December 1994.

[Lio94]   J. J. Liou, L. L. Liou and C. I. Huang, "Analytical Model for the AlGaAs/GaAs Multiemitter Finger HBT Including Self-Heating and Thermal Coupling Effects," IEE Proc. Circuits, Devices and Systems, vol, 141, no. 6, pp. 469-475, December 1994.

[Lio93]   Lee L. Liou, John L. Ebel and Chen I. Huang, "Thermal Effects on the Characteristics of AlGaAs/GaAs Heterojunction Bipolar Transistors Using Two-Dimensional Numerical Simulation," IEEE Transactions on Electron Devices, vol. 40, no. 1, pp. 35-43, January 1993

[Lio96]   L. L. Liou, B. Bayraktaroglu and C. I. Huang, "Theoretical Thermal Runaway Analysis of Heterojunction Bipolar Transistors: Junction Temperature Rise Threshold," Solid-State Electronics, vol. 39, no. 1, pp. 165-172, January 1996.

[Liu93]   William Liu, Steve Nelson, Darrell G. Hill and Ali Khatibzadeh, "Current Gain Collapse in Microwave Multifinger Heterojunction Bipolar Transistors Operated at Very High Power Densities," IEEE Transactions on Electron Devices, vol. 40, no. 11, pp. 1917-1927, November 1993.

[Liu95a]   William Liu and Ayca Yuksel, "Measurement of Junction Temperature of an AlGaAs/GaAs Heterojunction Bipolar Transistor Operating at Large Power Densities," IEEE Transactions on Electron Devices, vol. 42, no. 2, pp. 358-360, February 1995.

[Liu95b]   William Liu, "The Interdependence Between the Collapse Phenomena and the Avalanche Breakdown in AlGaAs/GaAs Power Heterojunction Bipolar Transistors," IEEE Transactions on Electron Devices, vol. 42, no. 4, pp. 591-597, April 1995.

[Man90]   A. M. Mansanares, A. C. Bento, H. Vargas, N. F. Leite and L. C. M. Miranda, "Photoacoustic Measurement of the Thermal Properties of Two-Layer Systems," Physical Review B, vol. 42, no. 7, pp. 4477-4486, September 1990.

[Mar93]   Antoine Marty, Thierry Camps, Josiane Tasselli, David L. Pulfrey and Jean Pierre Bailbe, "A Self-Consistent DC-AC Two-Dimensional Electrothermal Model for GaAlAs/GaAs Microwave Power HBT's," IEEE Transactions on Electron Devices, vol. 40, no. 7, pp. 1202-1210, July 1993.

[McA92]   Colin C. McAndrew, "A Complete and Consistent Electrical/Thermal HBT Model," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 200-203, 1992.

[Moi94]   Shahriar Moinian, Peter Feldmann and Robert C. Melville, "Chip-Level Electro-Thermal Simulation of Bipolar Transistor Circuits," in Proceedings IEEE Bipolar/BiCMOS Circuits and Technology Meeting, Minneapolis, MN, pp. 123-126, 1994.

[Mue64]   O. Mueller, "Internal Thermal Feedback in Four-Poles Especially in Transistors," Proceedings of the IEEE, vol. 52, pp. 924-930, August 1964.

[Mul77]   Richard S. Muller and Theodore I. Kamins, Device Electronics for Integrated Circuits, John Wiley & Sons, Inc., New York, 1977.

[Nag75]   L. W. Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits," Electron. Res. Lab. Memo. ERL-M520, University of California, Berkeley, 1975.

[Nak95]   Tohru Nakamura and Hirotaka Nishizawa, "Recent Progress in Bipolar Transistor Technology," IEEE Transactions on Electron Devices, vol. 42, no. 3, pp. 390-398, March 1995.

[Nis91]   H. Nishizawa, S. Azuma, T. Yoshitake, K. Yamada, T. Ikeda, H. Masuda and A. Anzai, "A Fully SiO_2-Isolated Self-Aligned SOI-Bipolar Transistor for VLSIs," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 53-58, 1991.

[Ona95]   Takahiro Onai, Eiji Ohue, Yohji Idei, Masamichi Tanabe, Hiromi Shimamoto, Katsuyoshi Washio and Tohru Nakamura, "Self-Aligned Complementary Bipolar Technology for Low-Power Dissipation and Ultra-High-Speed LSI's," IEEE Transactions on Electron Devices, vol. 42, no. 3, pp. 413-417, March 1995.

[Ozi93]   M. Necati Ozisik, Heat Conduction, Second Edition, John Wiley & Sons, Inc., New York, 1993.

[Pil94]   Lawrence Pillage and Ronald A. Rohrer, "The Essence of AWE," IEEE Circuits & Systems Magazine, pp. 12-19, September 1994.

[Pru94]   A. Pruijmboom, C. E. Timmering and J. J. E. M. Hageraats, "18 ps ECL-Gate Delay in Laterally Scaled 30 GHz Bipolar Transistors," in Technical Digest IEEE International Electron Devices Meeting, San Francisco, CA, pp. 825-828, 1994.

[Sab95]   SaberGuide and SaberScope Manual, Version 4.0, Analogy, Inc., Beaverton, OR, December 1995.

[Sch84]   A. Schutz, S. Selberherr and H. W. Potzl, "Temperature Distribution and Power Dissipation in MOSFETs," Solid-State Electronics, vol. 27, no. 4, pp. 394-395, April 1994.

[Sei93]   Ulrich Seiler, Eric Koenig, Peter Narozny and Heinrich Dambkes, "Thermally Triggered Collapse of Collector Current in Power Heterojunction Bipolar Transistors," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 257-260, 1993.

[Sha83]   D. K. Sharma and K. V. Ramanathan, "Modeling Thermal Effects on MOS I-V Characteristics," IEEE Electron Device Letters, vol. EDL-4, no. 10, pp. 362-364, October 1983.

[Shi96]   Takeo Shiba, Yoichi Tamaki, Takahiro Onai, Yukihiro Kiyota, Tokuo Kure and Tohru Nakamura, "A Very Small Bipolar Transistor Technology with Sidewall Polycide Base Electrode for ECL-CMOS LSI's," IEEE Transactions on Electron Devices, vol. 43, no. 9, pp. 1357-1363, September 1996.

[Shu90]   Michael Shur, Physics of Semiconductor Devices, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1990.

[Sze88]   Vladimir Szekely and Tran Van Bien, "Fine Structure of Heat Flow Path in Semiconductor Devices: A Measurement and Identification Method," Solid-State Electronics, vol. 31, no. 9, pp. 1363-1368, September 1988.

[Ten95]   B. M. Tenbroek, W. Redman-White, M. S. L. Lee and M. J. Uren, "Comparison of SOI MOSFET Self-Heating Measurements by Gate Resistance Thermometry and Small-Signal Drain Admittance Extraction," in Proceedings IEEE International SOI Conference, Tucson, AZ, pp. 48-49, October 1995.

[Ten96]   Bernard M. Tenbroek, William Redman-White, Michael S. L. Lee, R. John T. Bunyan, Michael Uren and Kevin M. Brunson, "Characterization of Layout Dependent Thermal Coupling in SOI CMOS Current Mirrors," IEEE Transactions on Electron Devices, vol. 43, no. 12, pp. 2227-2232, December 1996.

[Tu94]   Robert Tu, Clement Wann, Joseph King, Ping Ko and Chenming Hu, "SOI MOSFET Modeling Using an AC Conductance Technique to Determine Heating," in Proceedings IEEE International SOI Conference, Nantucket Island, MA, pp. 21-22, October 1994.

[Web91]   L. Weber and E. Gmelin, "Transport Properties of Silicon," Applied Physics A, 53, pp. 136-140, 1991.

[Win67]   Richard H. Winkler, "Thermal Properties of High-Power Transistors," IEEE Transactions on Electron Devices, vol. ED-14, no. 5, pp. 260-264, May 1967.

[Wol90]   Stanley Wolf, Silicon Processing for the VLSI Era, Volume 2: Process Integration, Lattice Press, Sunset Beach, California, 1990.

[Wor97]   Glenn Workman, Jerry Fossum, Srinath Krishnan and Mario Pelella, "Physical Modeling of Temperature Dependences of SOI CMOS Devices and Circuits Including Self-Heating," submitted to IEEE Transactions on Electron Devices.

[Yam93]   Tad Yamaguchi, Sudarsan Uppili, Galen Kawamoto, June Lee and Shaun Simpkins, "Process and Device Optimization of A 30-GHz fT Submicrometer Double Poly-Si Bipolar Technology," in Proceedings IEEE Bipolar Circuits and Technology Meeting, Minneapolis, MN, pp. 136-139, 1993.

[Zwe95a]   D. T. Zweidinger, J. S. Brodsky and R. M. Fox, "A Physical Thermal Resistance Model for Vertical BJTs on SOI," in Proceedings IEEE International SOI Conference, Tucson, AZ, pp. 84-85, October 1995.

[Zwe95b]   D. T. Zweidinger, R. M. Fox, J. S. Brodsky, T. Jung and S. -G. Lee, "Extraction of Thermal Parameters for Bipolar Circuit Simulation," in Proceedings IEEE Bipolar/BiCMOS Circuits and Technology Meeting, Minneapolis, MN, pp. 78-81, October 1995.

[Zwe96]   David T. Zweidinger, Robert M. Fox, Jonathan S. Brodsky, Taewon Jung and Sang-Gug Lee, "Thermal Impedance Extraction for Bipolar Transistors," IEEE Transactions on Electron Devices, vol. 43, no. 2, pp. 342-346, February 1996.

[Zwe97]   David T. Zweidinger, Modeling of Transistor Self-Heating for Circuit Simulation, Ph. D. Dissertation, University of Florida, August 1997.

BIOGRAPHICAL SKETCH

Jonathan Brodsky was born in Philadelphia, Pennsylvania, on August 3, 1969. In 1991, he received the Bachelor of Science degree in electrical engineering from Lafayette College, Easton, Pennsylvania. He received the Master of Science degree from the University of Florida, Gainesville, in 1993, and in the same year began working towards the Doctor of Philosophy degree in electrical engineering.

During the summer of 1995, he worked at the Motorola Advanced Products Research and Development Laboratory, Austin, Texas, where he helped characterize self-heating in SOI MOSFET's. Upon graduation he will join Texas Instruments, Inc. in Dallas, Texas, as a member of the Mixed Signal Products group working in the area of ESD protection.

He is a member of The Institute of Electrical and Electronics Engineers.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Fox, Chairman
Associate Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark E. Law
Professor of Electrical and Computer
Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William R. Eisenstadt
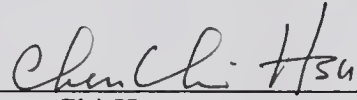Associate Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John G. Harris
Assistant Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Chen-Chi Hsu
Professor of Aerospace Engineering,
Mechanics and Engineering Science

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1997

Winfred M. Phillips
Dean, College of Engineering

Karen A. Holbrook
Dean, Graduate School